THE MANCHESTER SCHOOL OF ECONOMIC AND SOCIAL STUDIES

VOLUME XXV

1957

Reprinted with the permission of the original publishers
KRAUS REPRINT LTD.
Nendeln, Liechtenstein
1967

Printed in Germany

Lessing-Druckerei — Wiesbaden

CONTENTS

		Page
No. 1. January, 1957.		
EXPENDITURE, CAPITAL GAINS AND THE BASIS OF PROGRESSIVE TAXATION	WILLIAM VICKREY	1
THE NORTHERN WEAVERS, INDEPENDENT LABOUR REPRESENTATION AND CLITHEROE, 1902 ..	FRANK BEALEY	26
GAMES, CAPITALISM AND GENERAL LOCATION THEORY	M. L. GREENHUT	61
THE TAXATION OF OVERSEAS PROFITS	PAUL STREETEN	89
No. 2. May, 1957.		
THE GROWTH OF LIFE ASSURANCE IN U.K. SINCE 1880	JACK JOHNSTON G. W. MURPHY	107
AN ENGEL-TYPE CURVE FOR CASH..	JOHN SPRAOS	183
PRIVATE COMPANY FINANCE IN THE POST-WAR PERIOD	F. R. JERVIS	190
No. 3. September, 1957.		
THE CASH REWARD FOR AN HOUR'S WORK UNDER THREE PIECE-WORK INCENTIVE SCHEMES ..	TOM LUPTON SHEILA CUNNISON	213
FACTOR ENDOWMENTS, INTERNATIONAL TRADE AND FACTOR PRICES	H. G. JOHNSON	270
THE LEGAL MINIMUM WAGE, EMPLOYERS AND TRADE UNIONS: AN EXPERIMENT	BRIAN MCCORMICK H. A. TURNER	284
A NOTE ON SECULAR SWINGS AROUND ALTERNATIVE TRENDS	ROSEMARY WALKER	317

Digitized by the Internet Archive
in 2024

https://archive.org/details/manchester-school_1957_25

Expenditure, Capital Gains and the Basis of Progressive Taxation

Although the idea of a progressively graduated tax based on aggregate personal expenditure is by no means new, the lack of practical experience with any such tax is a perennial excuse for speculation about what such a tax would involve in practice, and the appearance of Nicholas Kaldor's recent book [1] advocating such a tax as a substitute for the surtax on personal incomes, following close on a series of far-ranging Royal Commission reports [2] is a further occasion for a re-examination of this tax in the light of present day income tax practices.

I. Income and Expenditure Taxes Compared in Ideal Terms.

In this area, more than in most, there is a sharp distinction to be drawn between considerations that would apply to the comparison of theoretically perfect taxes of the various types and those that would apply to the taxes as actually administered. Kaldor's treatment suffers considerably from a tendency to compare an idealized expenditure tax with the income tax as it exists, which is inappropriate unless indeed the tacit assumption is made that by turning over a new leaf a degree of perfection can be attained that has become difficult or impossible in the older form of taxation by reason of the accumulation of special vested interests and commitments to untenable positions. Kaldor's lack of success in getting the majority of the recent Royal Commission to accept a proposal for fuller taxation of capital gains [3] might indeed thus partly justify his approach. On the other side, even less tolerable is the reverse error of ignoring the difficulties of the present system, merely because they are somehow being lived with, while

[1] Nicholas Kaldor, *An Expenditure Tax*, London (Allen and Unwin), 1955, pp. 249. 18/-.

[2] Royal Commission on the Taxation of Profits and Income: *First Report*, 13th Feb., 1953, Cmd. 8761; *Second Report*, 9th April, 1954, Cmd. 9105; *Final Report*, 20th May, 1955, Cmd. 9474.

[3] See Cmd. 9474, pars. 80-117, pp. 25-40; pars. 34-84, pp. 365-382.

making the most of all of the difficulties, real or imagined, associated with the new.

Even-handed comparison of two taxes in terms of imperfect practice is difficult, however, not only because imperfection is always more complex than perfection, but because the nature of the imperfections to be encountered in practice are difficult to predict, particularly with respect to the new law. As a first approximation, therefore, it is helpful to compare a reasonably perfected income tax with an equally idealized expenditure tax ; when this has been done it will be easier to assess what changing to a new basis of taxation may have to offer as compared to reform of the existing tax.

If we begin accordingly by considering on the one hand a comprehensive income tax that insures that all forms of income, whether as dividends, interest, or capital gains sooner or later pay income tax on the same footing, and in which irregularities in the receipt or realization of income are adequately dealt with by some system of averaging, and on the other a correspondingly perfected spendings tax, we find that at least some of the differences between the two forms of taxation disappear.

Eliminating Bias against Risk.

For example, it is quite true that income tax laws as actually applied in nearly all countries have a considerable bias against risky investment.[1] This arises in some cases because with annual assessment and inadequate averaging under progressive rates the realization of a large profit will push income up into the higher surtax brackets, whereas a corresponding loss, even if fully deductible, will cause tax abatement at a lower marginal rate. In other cases, particularly the development of new products or processes by firms established for this purpose, profits, if the venture is successful, often present themselves most naturally in the form of dividends at a relatively high level, taxable at the full rates as income, whereas if the venture is a failure, so that much or all of the investment is lost, this will ordinarily rate as a capital loss, deductible, if at all, only subject to various restrictions and resulting in a relatively low level of tax abatement. A somewhat similar phenomenon

[1]See, for example, Kaldor, *op. cit.*, chapter III, pp. 102-129.

arises even more specifically if the investment is in bonds subject to risk of default of principal, in which case again the higher rate of interest obtained in consideration of the risk is fully taxed, but the loss in the event of default gets inadequate tax consideration.

But neither of these effects is inherent in the concept of an income tax. With full consideration of all gains and losses in the tax base, averaging of income over a sufficiently long period would greatly reduce the difference between the rate at which a gain would increase tax and the rate at which a loss would diminish it, since the marginal rate would depend not on the income for a single year but on the income of an extended period taken as a whole, in relation to which the gain or loss would be much smaller and thus cause much less variation in the point on the scale of progression at which the taxpayer would find himself. Moreover, to the extent that this bias against risky investment persists even with adequate averaging, it may be considered to reflect a consistent social policy. For if the progression of the tax schedule is deemed a reflection of the extent to which public policy looks with disfavour on inequality in the distribution of income, then it will be equally in accord with public policy to discourage individuals to place such large parts of their capital in such hazardous ventures as would carry an unduly heavy risk of greatly enriching or impoverishing them. Such a degree of bias against risky investment is inherent in any progressive tax system, including an expenditure tax, since the eventual spending of the proceeds of a successful venture would be subject to higher rates of tax than would the expenditure foregone in case of loss.

Consistent treatment under almost any concept of income would seem to require the avoidance of situations where the revenue plays "heads I win, tails you lose" with the taxpayer, and the allowance of full deduction of losses in all cases where the profits, if realized, even if in a somewhat different nominal form, would have been taxed. But it is extremely difficult to do this consistently unless substantially all gains and losses are brought into consideration in the determination of net income (with the possible exception of certain losses not of an investment nature, such as gambling losses, or casualites related to

consumption such as a mishap to a personal automobile). Losses from ventures whose successful outcome will normally result in taxable profits appear in a wide variety of ways, so that it is difficult to deny deductibility to any wide category of loss from investment without introducing an anti-risk bias. And it is hardly to be expected that deductibility should be allowed for any given type of loss without correspondingly taxing the gains in what appear to be analogous cases. Thus a full taxation of capital gains would seem to be a prerequisite for avoiding this anti-risk bias in the income tax.

Curiously enough, it is as a means of mitigating this bias that partial or complete exemption of capital gains is often advocated. But it is doubtful whether any kind of capital gains exemption can do anything but substitute one kind of bias for another, namely bias in favour of investments that are likely to reap their return in capital gain form over those more likely to reap returns in forms taxable in full; this distinction does not necessarily involve any significant element of risk, witness the case of the gilt-edged short-term bond selling at a discount because its coupon rate has been fixed below the market rate, possibly with tax advantages in mind. Insofar as the elimination of bias against risk is concerned, averaging plus inclusion of capital gains and losses in the income tax base can be considered an alternative to shifting to an expenditure tax. Kaldor would appear to be in error in his repeated assertion that a tax on accrued income would discriminate against risk more severely than the existing income tax.[1]

Offsetting Pressures for Retention of Corporate Earnings.

Another point where existing income taxes offend is in their impact on corporate finance and more particularly on the disposition of corporate profits. Especially where corporations are owned predominantly by shareholders in the upper surtax brackets, the tax results in considerable pressure for companies to avoid the distribution of taxable dividends, and instead to retain large fractions of their net earnings, either for expansion of the company's own business, retirement of debt, or investment in securities of third parties. There is considerable difference of

[1] *Ibid.*, pp. 117,121.

opinion as to whether this is on balance desirable, and the balance of considerations may shift from one side to the other over time; but whatever position is taken in general, the specific form in which the tax pressure is exerted seems to bear no relation to any public purpose. From one point of view such retention is undesirable in that it subjects the investment process to biases in the direction of pandering to the already suspect desire of corporation executives to add to their economic imperium in possible disregard of productivity considerations; in that it reduces the supply of investment funds available to rapidly growing smaller corporations through the flotation of new securities (as a supplement to what they could secure through the retention of their own earnings where this is justified by prospective returns); and in that it tends to encourage the growth of established companies with consequent tendencies for monopoly and the dominance of the economy by the very large companies. In the United States, concerns of this nature have been reflected in such phenomena as the short-lived surtax on undistributed profits of 1936-39, the notorious section 102 surtax on corporations having improper accumulations of surplus, and the special penalty taxes on personal holding companies that accumulate undistributed profits.

In Britain, the quite properly greater concern for anti-inflationary considerations and the greater tolerance towards monopoly have resulted in what is to the American observer the amazing spectacle of a profits tax which discriminates against distributed profits, adding its pressures to those of the surtax. Yet it is highly doubtful whether such a provision has had any important direct anti-inflationary effects (particularly as the individuals whose dividends were thereby kept low probably had low propensities to consume and were in many cases in a position to finance consumption expenditures in other ways), even though such measures may have filled a need in making restrictive policies in other directions more palatable politically. Granting everything that may be said in favour of internal financing as being economical in avoiding the expenses connected with the flotation of securities, and as directing funds more or less automatically in the directions that have been proved most profitable in the past, nevertheless it seems hardly

appropriate to further reinforce these factors, which are quite adequately under consideration by company directors, by artificial and irrelevantly modulated tax pressures. Moreover the consequence for the present seems to be that such a provision, once enacted, is difficult to eliminate without causing a windfall to shareholders in the form of increased share prices, which in turn could only be tolerated on equity grounds if there were adequate provision for the taxing of the resulting capital gains.

It is entirely possible to design an income tax structure that is reasonably free from such irrelevant and unintended influences on company policy. Indeed, under an income tax with full inclusion of gains and losses and cumulative averaging in which the interest factor is taken into account,[1] the deferment of the realization of the income by the shareholder that results from the retention of earnings by the company would not affect the ultimate burden of the shareholder, and such a tax would therefore be reasonably neutral on this point. Even if a less pretentious form of averaging is used that does not take into account the interest factor, it would be possible to keep the tax system as a whole reasonably neutral by applying to the accumulated undistributed earnings of companies an annual tax at say 1% or 2%, this being deemed to be equal to an interest charge of say 5% on the surtax which is being postponed by reason of the non-distribution of the earnings.[2] Problems of discrimination against risk and pressures on dividend policy are thus both capable of being taken care of within the framework of the income tax concept, and it is accordingly only to a limited extent that an expenditure tax can be considered necessary to achieve these ends.

Comparative Effects on Saving : the Net Rate of Return.

A third major claim for the expenditure tax is that it will encourage saving. How weighty this consideration is will depend on the context of general economic conditions, anticipations regarding the future of the tax, the motives that

[1]Cf. the writer's *Agenda for Progressive Taxation*, (Ronald, New York, 1947), Chapter V, esp. pp. 172 ff.

[2]Cf. the writer's "A Reasonable Undistributed Profits Tax," *Taxes*, Feb., 1945, pp. 122-127.

influence the saving of individuals, and possible criteria as to the socially desirable amount of saving. If the tax is thought of as a permanent and long established tax, then as Kaldor and others before him have shown,[1] the effect of changing to an expenditure tax is to permit the individual to adjust the time pattern of his consumption, through corresponding adjustments in his savings program, so that it will bear a correct relation to the relative social costs of providing goods and services at different times as this relative cost is reflected in the market rate of interest, whereas the income tax distorts this relationship, making present rather than future consumption unduly attractive to the individual in spite of its higher drain on the resources of the community. This is the traditional analytic argument for preferring the expenditure tax to the income tax, and there is certainly no denying its validity, nor any possibility of amending the income tax so as to avoid this defect without in effect converting it into an expenditure tax. Opinions will differ, however, as to how important the changes that would occur on this basis would be. It is at least arguable that the increase in the savings of an individual in a 16 shilling in the pound tax bracket that results from increasing the net rate of return on his savings, after tax, from say 1% to 5% by shifting from an income tax to an expenditure stax, and similarly for other taxpayers, is not of paramount importance to the economy as a whole.

Savings in an Emergency.

There is a case of a different order to be made, not discussed by Kaldor, for a temporary expenditures tax to be applied at a time of great general shortage of consumer goods, as during a war or other emergency period.[2] In this case the tax is specifically not expected to continue after the emergency is over, or at least only at much lower rates, so that deferment of expenditure not only defers the tax, but reduces the applicable rate, and the incentive for individuals to save, at least temporarily, is greatly enhanced. In such contexts the

[1]Kaldor, *op. cit.*, Chapter II, pp. 79-101.

[2]Cf. the writer's, "The Spendings Tax in Peace and War," 43 *Columbia Law Review*, pp. 165-170 (1943).

argument for the expenditure tax is much stronger than it is for a permanent tax ; indeed it can be considered as a form of graduated generalized rationing. One could even argue that the possibility of such emergency use would be a reason for using the tax in normal times in order to have the administrative machinery ready for the emergency.

Effects on the Concentration of Economic Power.

A much more controversial issue that arises in considering an expenditure tax as a permanent tax is the effect of such a change on the concentration of the ownership of wealth and of economic power. Kaldor claims that "by making the expenditure tax sufficiently progressive it would always be possible to prevent the rich from saving too much," [1] but this is patently erroneous, since spendings tax rates, however progressive, could have no effect on a man who persisted in living on a frugal scale in spite of his great wealth, even leaving out of account the easing of this frugality by the non-taxable perquisites that almost invariably accompany the possession of great wealth. Possibly there may be some who would view such accumulations of wealth in private hands with equanimity and even approval so long as the spendings tax rates were such as to induce their possessors to treat them as funds held in trust rather than for personal gratification through consumption of the income. It can be hazarded, however, that the weight of opinion would regard with some misgiving even that accumulation that might be made by a man such as Henry Ford within his own lifetime, let alone an accumulation continued unchecked over several generations. If a spendings tax is to be substituted for an income tax, some complementary form of progressive taxation seems called for to check the more extreme accumulations.

Corresponding Adjustment of Death Duties.

The usual recourse at this point is to invoke heavy death duties as the remedy. Even without balking at the very substantial possibilities that would still remain for the accumulation of wealth within a single lifetime, one can point

[1] Kaldor, *op. cit.*, p. 97.

to the very serious difficulties, often overlooked, that stand in the way of setting up an effective, sharply progressive succession tax that is reasonably free from untoward consequences. Present-day death duties are subject to major avoidance through gifts *inter vivos*, through skipping generations by the setting up of trusts or otherwise, and in many other ways ; if these avenues of avoidance were effectively blocked, another very troublesome one of international transfers and migrations would arise that could be blocked only at the expense of seriously interfering with legitimate and desirable international movements of capital ; the present international convention to the effect that only the jurisdiction of situs may tax real property would have to give way. Kaldor suggests that an appropriate succession duty should vary the rate of tax according to the net wealth of the recipient ; [1] possibly also the difference in age between transferor and transferee would need to be taken into consideration [2] so as to avoid the incentives that would otherwise develop if the rates were at all steep to skip over as many generations as the rules against perpetuities will allow, with consequent undesirable tying up of property in various kinds of trusts to the detriment of the supply of venture capital and of flexibility in the use of property.

Thus at least one of the liabilities of the expenditures tax is that its substitution for an income tax renders more crucial the solution of the far from easy problem of developing an adequate succession duty. To make a fair comparison between the two taxes thus requires considering not merely the change to the expenditure tax itself, but also the appropriate changes in succession taxation that would be called for. This necessarily greatly complicates the comparison. However, a not too difficult comparison becomes available if in lieu of imposing a specific succession tax we merely on the one hand apply the expenditures tax to all gifts and bequests, considering them to be a form of personal consumption of the donor, while under the income tax on the other hand we consider gifts and bequests to be taxable income to the recipient. Such an arrangement would not be satisfactory in practice, because of the avoidance

[1] *Ibid.*, p. 101.

[2] Vickrey, *Agenda for Progressive Taxation*, pp. 216 ff.

possibilities mentioned above, so that caution must be used in interpreting the comparison. Assume a stable price level, and suppose further that in both the income tax and the expenditure tax, to avoid the effects of fluctuations from year to year in the tax base, which with the inclusion of gifts and bequests would be particularly severe, the tax is assessed on a cumulative basis, comparable to the method used for P.A.Y.E., the final tax liability being assessed on the basis of an aggregate of the income or expenditure over the entire period, prior tax payments being considered a deposit against the ultimate liability, compound interest being credited at an appropriate rate on this deposit.[1] The aggregate amount brought to account for tax purposes for any individual will be the same in both cases, the difference between the impact of the two taxes being entirely due on the one hand to the relative weights given the quantities reported for the various years in aggregating them, and on the other hand to the timing of the interim payments on account.

Weighting Effects.

As to weighting, in the case of an income tax it would be appropriate to make each year's income count the same as any other, to avoid any arbitrary effects on the tax burden due to shifts in the time of realization of income ; for the spendings tax it would be more appropriate to use a present value of the expenditures of various dates at compound interest so as properly to express the slighter burden imposed on society by the later consumptions. The weighting is then such as still to preserve the difference between the two taxes in the net rate of return offered in terms of future consumption or ability to make bequests in return for present abstinence : under the expenditure tax this remains equal to the market rate of interest, in the income tax it is reduced by the marginal rate of tax, and thus the increased incentive to save produced bythe spendings tax can be represented by this change in interest rate.

Timing Effects.

As to the difference in the timing of payments for the two taxes, this depends on the way in which income is realized for tax purposes. At one extreme, if income is realized for tax

[1] *Loc. cit.*, note 5.

purposes only as needed to provide funds to finance personal expenditure, then the timing of the payments of the two taxes becomes quite similar. This is indeed the ultimate extreme of the policy of allowing accelerated depreciation, or granting initial capital allowances, in that in effect all investments are written off to zero as soon as made, with the implication that all subsequent amounts realized from such assets, of whatever character, would then be treated as income and not return of capital. The writing down process would presumably be limited in the case where the taxpayer had indebtedness, so as not to produce a negative nominal net worth. The effect of shifting from such a tax to an expenditure tax would then be limited to the incentive effect produced by the fact that future expenditure would be taxed relatively more lightly than present expenditure, and there would not be any element of leaving the more frugal taxpayer with greater resources at his command under the expenditures tax than under such an income tax.

Another way of looking at the matter is to consider that allowing the taxpayer to postpone the realization of income is in effect to allow him to postpone the payment of the corresponding tax, which can be considered equivalent to a loan of the corresponding funds from the government to the taxpayer. If, as under an income tax, this postponement will often depend on purely nominal differences in financial arrangements that are more or less under the control of the taxpayer, it becomes necessary, in the interests of equity between taxpayers differing in the flexibility of their finances and of avoiding the distortion of normal financial arrangements in response to tax pressures, to charge an appropriate rate of interest for these deferments or loans. This is not in present practice done, at least not in any systematic way, but it can readily be done in conjunction with cumulative averaging. On the other hand if the deferment of tax is co-ordinated with a deferment of consumption, as under an expenditures tax, then one can without inequity omit the charging of interest on the tax deferred, thus producing the appropriate incentive for saving. One can readily afford to provide such incentives for saving when the saving is net aggregate saving, but can generally ill afford to provide even more moderate incentives for particular forms of saving when

these may be offset by dissaving in other forms so that the incentives would operate more to divert savings from its natural channels than to increase the total flow of savings.

Thus even within the income concept it would be possible to provide a procedure whereby tax on reinvested income would be deferred, thus avoiding the tendency of heavy income taxation to impinge on the funds that might be used by frugal and skillful entrepreneurs to expand the scale of their operations. What is inherently and necessarily different in the expenditure tax approach is merely that there is appropriate differentiation between the present values of the tax attracted by increments of present consumption as compared with increments of future consumption.

* * * * *

II. Comparison in Practical Terms.

Inadequate Succession Tax Adjustment.

If on the theoretical level the difference between the two principles of taxation comes down to this relatively minor basic element of weighting through time, what of the differences that arise on a more practical level? To begin with, it must be admitted that a shift from an income tax to an expenditure tax is unlikely in practice to be coupled with a corresponding shift in the succession taxes of the exact nature assumed in the above comparisons. Such a shift would indeed involve new problems of administration and avoidance practices, and on the whole any shift in succession taxes is likely to be smaller than that called for to achieve the theoretical balance relative to the *status quo ante,* so that the expenditure tax alternative is likely to involve rather more of a problem of increased concentration of economic power than the idealized version might indicate; on the other hand, the stimulus to savings may at the same time be correspondingly greater.

Effects of Lack of Averaging.

Secondly, while there seem to be no insuperable difficulties to the adoption of some form of long-term averaging, either for income tax or expenditure tax purposes, there is little

relevant experience to go on, and such averaging might fail of adoption. In the absence of such averaging the incentive effects of the two taxes become rather more uncertain, as they come to depend considerably on anticipations held concerning future rates of tax as compared to present rates. Again, without averaging, accelerated writing off under an income tax is likely to produce serious inequities if carried beyond a very restricted range, unless indeed it is so generalized as to become tantamount to a shift to an expenditures tax.

Introduction of an expenditure tax might have rather unfortunate effects if there is no averaging to smooth the impact, particularly if it is introduced either gradually over a period of years, or with considerable advance discussion rather than as a full-scale *fait accompli* (which would be difficult enough under British legislative procedures; it would be virtually impossible in the United States). Taxpayers would be on fairly definite notice, possibly for a number of years, that with high probability spendings tax rates would be on the increase, so that there would be considerable tax advantage in advancing expenditures from the later to the earlier years of lower rates, or possibly to periods before the effective date of the tax. The "buy now before Budget" appeal would no longer be limited to a narrow range of goods or to the offer of better hire-purchase terms, but be universalized in terms of net expenditure, so that for a period the effect of the new tax on savings might actually be adverse. Kaldor seems to feel that this difficulty would be minimized by natural limits to the short-run supply of consumer durables [1] but if the number of taxpayers to be brought under the tax at any particular stage is small, this limit would seem not very restrictive, particularly since if rates are at all steep, other forms of expenditure such as travel may well be capable of being pushed ahead. Short of springing the tax as a surprise on a full scale, which seems politically unlikely, such effects could be avoided only by advancing the effective date of the new tax sufficiently so that taxpayers by and large would have had little advance warning, and using a cumulative method of assessment from that date. Even this might be difficult if, as seems otherwise desirable,

[1] Kaldor, *op. cit.*, p. 220.

the gradualness of the introduction consists in part of starting with the top of the income distribution and gradually lowering exemptions so as to bring more and more taxpayers within the scope of the tax.

Even aside from this transition problem, averaging seems likely to play a more crucial role in a spendings tax than in an income tax. While it may well be that the degree of variation from year to year in the conceptual expenditure tax base of a given taxpayer would vary less from one taxpayer to another than that of the income tax base, the nature of the graduation that might be thought appropriate for an expenditures tax might well make these more important. And while in principle the expenditure reckoned with respect to consumer durables may well be smoothed out in time by considering only notional interest and amortization figures spread over several years rather than the initial outlay in a lump sum in the year of purchase, this would involve serious administrative difficulties were it extended to more than a very few items, such as housing and perhaps automobiles. Many types of "lumpy" expenditure, such as for travel, fail to lend themselves easily to the consumer durables treatment, and even where the consumer durables concept applies, as for example to many of the outlays involved in setting up a new home, the items are too numerous to make anything but some kind of aggregate treatment feasible. On the whole it would seem that if the expenditure tax is to achieve its objects it is highly desirable that it should be accompanied by some kind of averaging. Once we admit this, however, we must also admit that a similar application of averaging to the income tax and surtax would go far toward achieving many of the ends for which the shift to the expenditure tax is advocated, such as the mitigation of bias against risky investment.

Separate Taxation of Small Incomes.

But even more important than the question of averaging is the matter of how far can or should the practical application of the tax go in attempting to make the actual tax base conform to a theoretically consistent concept. Here it is high time that we gave much more extensive and explicit recognition to

the fact that income taxes, which in their initial stages were taxes imposed on a relatively small number of taxpayers at rates such that exact niceties in the determination of the base were not essential, have now become taxes which extend on the one hand to large masses of taxpayers where more than a very limited effort per taxpayer expended in the determination of the base would be wasteful of resources, and on the other to top marginal rates where defects in the base result in serious inequities and unwholesome influences on economic behaviour. This is of course recognized everywhere to the extent that administration is concentrated on the larger returns and there is thus a tendency to overtook minor items and errors in the smaller returns. In the United States this is further recognized in the simplified tables, the short form of return, and the optional standard allowance for deductions ; in France this finds expression in the practice of allowing small taxpayers to compute their tax on the basis of more or less arbitrary schedules or "forfaits" ; and in the United Kingdom of course, there is the separation of the assessment of the surtax from that of the income tax, though the practical effect of this separation is minimized by the assessment of both taxes on substantially the same base.

The requirements of these two segments of progressive taxation differ so greatly, indeed, as to raise the question of whether it is not in order to complete the separation and define the tax bases separately for the two in terms appropriate to each, rather than continue what can at best be an unsatisfactory compromise. The mass tax would thus be defined in terms facilitating collection at source, simple procedures and returns, and a minimum of details or minor adjustments. The progressive frosting would be defined independently in a manner such as to minimize opportunities for avoidance, untoward economic influences and arbitrary discriminations, and with ease of administration considered in relation to the number of taxpayers involved and the types of problems that are prevalent in dealing with such taxpayers.

Kaldor's proposal for an expenditure tax as a substitute for the surtax is indeed one way of effecting such a separation. But such a separation is also possible within the income tax

area. One might, for example, consider at this point that the recommendation of Kaldor and others in the dissenting minority of the Royal Commission that capital gains be subject to income tax but not to surtax [1] might well be inverted and turned into a recommendation that such gains and losses should be ignored for normal tax purposes but made subject to surtax, after grossing up or the equivalent. Such a treatment would be more in accord with a general principle of confining complications to that part of the tax applicable to the top bracket taxpayers, and at the same time would go much further than the minority suggestion in maintaining the progression of the tax and minimizing uneconomic influences flowing from the tax. Indeed, such a treatment could well be considered correct in theory insofar as capital gains in securities are a reflection of undistributed company earnings, since in such cases they will already have borne tax at the normal rate, and can be considered as equivalent to a net dividend.

Co-ordination of Mass Tax and Surtax.

As a minor technical matter, it should be observed at this point that if the base for a surtax differs from that for a normal tax, then in order to minimize the effect among taxpayers at the top of the scale of imperfections in the base of the normal tax, it is essential to adjust the base for the surtax appropriately. This can be done either by grossing up all income subject to surtax by including a notional normal tax even though some items of income are not subject to an explicit normal tax, or by using as a base for the surtax the income net of normal tax in all cases, with suitable adjustment in the surtax rates to be applied. For example, if the taxable income is the same for both taxes, it makes little difference whether we have a normal tax of 9/- in the £ and a surtax of 10/- on the grossed up income, or the same normal tax with a surtax at a rate of 10/11ths on the income net of normal tax. If however for reasons of administrative convenience or otherwise some items of income escape normal tax, then if these items are not grossed up but the surtax is nevertheless on a gross basis, these items will be taxed only at 50% instead of at 95%, whereas if they are

[1] Cmd. 9474, par. 62, p. 374.

grossed up, or if the surtax is assessed at the higher nominal rate on a net basis, they will bear tax at approximately 91%, as compared with the full combined rate of 95%. As the grossing up of items to include a notional normal tax where no such tax has actually been paid will be a little mysterious to most taxpayers, probably the preferable procedure is to assess the surtax at the higher rates on a base net of normal tax. This device might also serve to facilitate, psychologically, a partial retreat from excessively high top bracket rates.

The minority suggestion of taxing capital gains only under the normal tax does have the advantage of avoiding the need for averaging, since such gains would be taxed almost entirely at a flat rate, relatively few taxpayers to whom capital gains are significant being in the reduced rate income levels. But as averaging in any case seems quite an appropriate element of a surtax conceived of as a refined levy applied to relatively few taxpayers, this argument loses weight. Averaging, indeed, can serve as a means whereby the scope of the surtax might be acceptably narrowed : for example the surtax might be started with a rate of 3/- in the £ on income above £2,500, instead of at 2/- in the £ on income above £2,000, thus reducing the number of persons subject to surtax by about 25% ; [1] without averaging this sharp jump in the rate would tend to increase too severely the inequities resulting from fluctuating incomes, whereas with averaging this obstacle vanishes. On the incentive side, too, averaging moderates the jump, since at the time of deciding whether to work harder or not the taxpayer will ordinarily be somewhat uncertain as to whether his cumulated income will ultimately average out below or above the margin where the extra 3/- applies.

Choice of Base for a Mass Tax.

If we look first at the mass tax, it seems that the use of an expenditure concept here as the main basis of the tax would encounter considerable difficulty. To be sure, the existing normal tax is in effect something of a hybrid, containing several elements related to expenditure mingled with the income items,

[1] If rates above £3,000 remain unchanged, ranging from 3/6 to 10/- in the £, the loss in revenue would be about £10 million, out of a total income tax and surtax of about £2,000 million.

such as the allowances in respect of certain types of saving, to wit life insurance and building and loan societies, and on the other hand the inclusion in the tax base of certain periodic receipts which are in part a return of capital, such as some types of annuities. We have then not so much a question of proceeding from an existing income tax to a new expenditures tax, but a finding that approximation or compromise most suitably balancing economic and administrative considerations. One could, indeed, go a long way toward broadening the range of investments that qualify for allowances and extending the categories of realizations that attract tax. But it does not seem that one can do this to the extent that would approach a spendings tax concept without requiring the taxpayer to file a moderately detailed return, which for this level of taxation would seem to be something to be avoided if possible.

In the other direction, a somewhat closer approach to a consistent income concept would seem to be both desirable and attainable without serious added administrative burden. The allowance for life insurance, for example, so long as it is not generalized to a wide area of savings, would seem to be an unwarranted concession to a particular form of saving and might well be eliminated, at least as to new policies ; on the other hand allowance might well be made for the return of capital element in annuity payments[1] as also for the income element in lump sum benefit payments. On the other hand, as suggested above, all capital gains on transactions in domestic securities might be ignored regardless of the frequency of the transactions or their relation to the main business of the taxpayer, on the ground that the tax has already been borne at the source.

Information Requirements for Capital Gains and for Expenditure Taxation.

When it comes to considering what form of tax on incomes will be appropriate at the upper end, it is clear that no reasonably close approach to a consistent income concept can be obtained without taking capital gains and losses fully into

[1]Cf. Committee on the Taxation of Trading Profits (the "Tucker Committee"). *Second Report*, Cmd. 9063, pars. 477-505.

account. A strict accounting for such gains and losses, however, requires the eventual production of data on the cost of nearly all investments and on the proceeds of nearly all disinvestments, the only exceptions being those where cost and proceeds are identical, such as bank deposits. But this information comes very close to being sufficient to determine the base for an expenditure tax as well, since this is to be computed, not by adding up expenditure items, but rather by taking total receipts, deducting outlays for other than personal expenditure, and adjusting for changes in cash balances and bank deposits. Income items and the proceeds of disinvestments would constitute the great bulk of the receipts, the main other items being gifts, bequests, prizes, and the like which would not be too much trouble to keep track of; the principal deductible outlays would be investments, the figures for which would be the same as those on the cost side of capital gain calculations. For persons in the higher income ranges, among whom capital transactions are likely to be of some importance, there is thus very little difference between the administrative burden involved in the assessment of a refined income tax that covers capital gains and that for an expenditures tax. This is particularly true if outlays for consumer durables are assessed on a cash outlay basis, hire-purchase payments, for example, being used as a base rather than the nominal cash price, with averaging being relied on to avoid inequities.

The administrative difficulties involved in an expenditures tax confined to upper bracket taxpayers are thus of somewhat the same order of magnitude as those involved in the taxation of capital gains and losses. One should not, however, take the difficulties encountered in the United States as an index of these difficulties, for there a major part of the difficulty arises precisely because capital gains are taxed on a different basis from other forms of income, and it becomes necessary not only to compute the various gains but to determine whether they are capital gains or ordinary income, a matter which has over the years absorbed vast quantities of legal talent. If such gains were treated on exactly the same footing as other profits, the determination of these niceties, frequently of a hair-splitting nature, would be unnecessary. The administrative burden

would be eased at other points as well: determination of whether a given item should be treated as a current expense or a capital outlay, for example, would be less crucial, since in the long run any reduction in present profits brought to account will show up as an increased capital gain. In conjunction with cumulative assessment, indeed, such changes in accounting practice, and others involving depreciation, and accruals of all kinds, will leave the ultimate total tax burden unaffected, so that issues of this kind can safely be left to the discretion of the taxpayer's own accountants, and the revenue can be relieved of the burden of inquiring into such matters. For schedule D taxpayers, indeed, the result might be a net simplification.

There is, however, one important source of difficulty in the assessment of capital gains that is absent under the expenditure tax: for a newly established tax regime, especially, the assessment of capital gains sooner or later requires either a valuation of all assets of taxable persons as of an inaugural date, if gains accrued under the prior regime are to be exempted, of if they are not to be so exempted, a determination of costs of acquisition of such assets, often long after the event with the consequent difficulty of obtaining adequate information. Even with a long established assessment of capital gains, the need for the determination of the costs involved in a particular acquisition may not become evident immediately, for example if the taxpayer is not yet favored with an income of surtax magnitude, so that when the determination does take place, often only at the time of assessment for the year covering the date of sale, it may partake of some of the difficulties of archeological research. For the expenditure tax, on the other hand, the costs that are relevant are always those incurred during the expenditure year, so that they are thus much more readily ascertainable.

If it be admitted then that at the upper end of the progressive scale a refined tax base of some kind is essential, and that for an income tax this implies inclusion of capital gains and losses in the base, then there would seem to be little to choose between income and expenditure on administrative grounds, at least for those classes of taxpayers where property incomes and dealings in capital assets are of importance. In

this area, accordingly, one can make the choice on the basis of economic effects, political appeal, or other grounds.

The Advantages of a Fresh Start.

One consideration of a tactical nature that may in practice be of some importance is that the shift to a new principle of taxation gives an opportunity for a fresh start on the detailed provisions, with less need for tender treatment of old commitments and vested interests. In the United States, for example, one of the advantages sometimes cited for an expenditure tax is that it would side-step the issue of exempt government securities (and also, formerly, of exempt government salaries). In Britain the matters of annuities, income covenants, and the like are examples. Another element of this kind is that the change might facilitate a fresh look at the reasonableness of the top rates, since the basis of graduation would be quite different. As Kaldor points out,[1] the natural and from many points of view preferable method of assessing the tax is to exclude the tax itself from the base; thus a tax on an inclusive base of 60%, for example, becomes a tax of 150% on the exclusive base; the rate corresponding to the top income tax rate of 19/- in the £ would be 1900%, or £19 of tax on each £1 of expenditure.

But granting that among the highly refined tax bases needed at the top of the progressive scale, the expenditure tax might be as easy to administer as the income tax, are there not other alternatives that are acceptable and less difficult? The line of least resistance, indeed, is to follow the majority of the Royal Commission and carry on with the existing system with minor revisions. But surely, in the light of what has been said above, and as the minority of the Royal Commission insists, continuation of the present tax situation is not a satisfactory solution. When one considers the manifold opportunities for avoidance that exist under the present law, together with the degree to which incentives for avoidance are increased by pushing rates up to 90% and even 95%[2] together

[1] Kaldor, *op. cit.*, pp. 232-238.

[2] For example, raising the rate from an aggregate of 18/- to 19/- increases the revenue obtainable from the income affected by less than 6%, but it doubles the amount of fully taxable income that the taxpayer would be willing to sacrifice in order to obtain a receipt in non-taxable form or in a form taxable at a given lower rate.

with the relatively small amount of revenue produced by these last increments in the rates even if the tax base could be assumed to be unaffected,[1] it seems highly likely that these rates have entered the area of negative returns and that the last increments neither yield any net revenue in the long run nor contribute materially to the real progression of the tax, while they do drastically magnify the inequities between taxpayers who differ in the degree to which the loopholes are readily available to them. If nothing can be done to improve the tax base, retention of these top rates can only be considered a form of economic sadism, by which the wealthy are penalized without any consequent benefit to the rest of the community. The *status quo*, or anything approaching it, is not an acceptable alternative.

An Annual Surtax based on Net Worth.

There is, however, another alternative meriting consideration, and that is the imposition of an annual tax at progressive rates on net fortunes, at rates ranging up to say 3% to 6%. At first glance it might seem that such a tax would be unacceptable because its administration on a reasonably refined basis appears more difficult than a correspondingly refined expenditure or income tax covering approximately the same number of taxpayers at the top of the distribution; not only would some sort of balance sheet have to be established for each such taxpayer, but if serious pressures for the postponement of realization of gains are to be avoided, some kind of fairly frequent valuation of assets will be required. But this is offset to a considerable extent by the fact that at given high levels of progression the pressures and opportunities for

[1]Figures for 1953-54, the latest available at time of writing, indicate that out of a total income tax and surtax revenue of £2,008 million, of which £134 million was surtax, only £36 million would have been lost by levelling off the surtax rates at a maximum of 5/- in the £, or 70% combined income tax and surtax, even assuming no abatement of avoidance practices as a result of the reduction in rates. For a levelling off at 7/-, or an 80% combined rate, only £10 million would have been lost, while the last increment from 90% to 95%, which again doubles the incentives to avoidance felt by taxpayers with over £100 million of income altogether, produced only £2.4 million gross. Figures are derived from the 98th Report of the Commissioners of Her Majesty's Inland Revenue for the year ended 31st March, 1955, Cmd. 9667, pp. 14, 78.

avoidance and the consequent need for accuracy in the determination of the tax base seem less than with an income tax. This can be shown most vividly by a somewhat exaggerated example as follows : suppose that the group with net fortunes of 10 million pounds or over have incomes equal to 7% of their wealth ; an annual tax on net wealth at a marginal rate of 8% would correspond in progressivity of burden to a marginal income tax rate of 114%, but would have nowhere near the disastrous incentive effects that such an income tax would have : the income tax would make it to the taxpayer's advantage actually to suppress his income,[1] whereas the net worth tax would merely tend to discourage saving and encourage dissaving or dissipation of capital in gifts and personal expenditure ; it would still be to the advantage of the taxpayer to derive as much income as he could in any particular period from such capital as he then had. Thus even though the base of the net fortunes tax would be inherently more difficult to define and determine than that of an income or spendings tax, greater imperfections in definition and determination could be tolerated at high levels of progression than for the other taxes. The net administrative difficulty of achieving a satisfactory standard might thus be smaller.

Nevertheless, to throw on a net fortunes tax the entire task of providing progression at the top end of the scale is probably not an acceptable solution, as the relief that would

[1]Marginal combined rates exceeding 100% do occur in practice, particularly in the case of new businesses, though since the impact may be spread over several years the combined effect is often not fully appreciated. If for example a new business is started 1st Sept., 1953, and makes up its books to 30th June of each year, then the assessment for the year 1953-4 will be the income allocable to the period to 5th April, or 43/60 of the income for the ten months to 30th June, 1954 ; at a 9 shilling standard rate, this is equivalent to about 6/5 in the £. The entire income for this first accounting period will again be included in the assessment for 1954-5 and also for 1955-6, at 9 shillings in the pound each time, or a total of £1/4/5 in the £. As long as the proprietors have enough income to be taxable at the full standard rate, any expenditure, however wasteful, that qualifies as a deduction in computing net income for this initial period, will be a profitable investment in terms of income tax reduction over these three assessment years. If the proprietors are in addition subject to surtax, the situation is of course even more extreme. (See *Report of the Committee on the Taxation of Trading Profits*, Cmd. 8189, April, 1951, pars. 19, 24, 27).

thus result for those whose incomes are "earned" or derived from non-capitalized sources would probably be considered excessive. At most one could regard the introduction of the net fortunes tax as making it possible to reduce the top rates of the income tax to levels where the inequities and distortions produced by the present form of tax would be more nearly tolerable. And while a net fortunes tax might be regarded as in some measure redressing inequalities resulting from the greater availability of various forms of avoidance to property incomes than to earned incomes, this correction would be an overall matter and would still leave substantial individual inequities: the availability of loopholes would still vary greatly from one form of property to another. Indeed in some cases even salaries can be converted into capital gains, as has been done on a considerable scale in the American movie industry through film stars obtaining compensation in the form of a share in the film produced; a somewhat similar phenomenon in Britain was dealt with by the Royal Commission under the heading "Post-cessation receipts."[1] Considerable improvement may thus be obtained with the aid of a net fortunes tax, but it does not in itself provide a complete answer to the problem of progression.

There is, indeed, no royal road to effective, equitable, and reasonably distortion-free progression of the degree represented by the nominal rates standing on the statute books of most industrialized countries. The major alternatives, aside from continuing to tolerate the absurdities and inequities of the *status quo* with only minor revisions are: (1) to retreat from the attempt to apply steep progression at the top of the scale, to whatever degree is necessary to reduce the inequities and distortions arising from the present tax base to tolerable levels, (2) to substitute a net fortunes tax for some of the graduation of the income tax, partly as an offset to some of the inequities, (3) to reform the income tax, at least as applied to upper bracket taxpayers, so as to provide full inclusion of capital gains and losses and some form of averaging, and (4) to adopt a spendings tax for upper bracket taxpayers. Of these only the last two offer a fundamentally satisfactory

[1]Cmd. 9474, Chapter 12, pp. 80-83.

and sound solution; both involve a considerable additional administrative burden of about the same order of magnitude, the balance between them probably depending on the details of the way each tax would be applied. The economic advantages of the expenditure tax are not negligible, though as compared with a reformed income tax they are considerably less striking than might at first be thought. If ever enough resolution is mustered to deal with the problems of progressive taxation in a thoroughgoing manner, the expenditure tax will merit serious consideration, and Mr. Kaldor's book will have provided a valuable challenge.

WILLIAM VICKREY.

Columbia University.

The Northern Weavers, Independent Labour Representation and Clitheroe, 1902[1]

On 1st August, 1902, David Shackleton of the Darwen Weavers was returned unopposed as the candidate of the Labour Representation Committee for the Clitheroe Division of Lancashire. This success, which sent the third independent labour M.P. to Westminster, was of the greatest significance. However, it has not received the attention from historians it deserves.[2] The fight of Will Crooks at Woolwich in March, 1903, as well as being more of a contemporary sensation, has interested posterity more than the earlier uncontested triumph in the remoteness of North-East Lancashire. Yet Clitheroe rather than Woolwich portends that upsurge in Labour strength which was in 1906 to sweep thirty Labour M.P's into the House of Commons ; for Clitheroe witnessed the first Labour victory since the Taff Vale judgment, with its new interpretation of the legal position of trade unions, had thrown the industrial world into confusion.[3]

The Labour Representation Committee was founded on 27th February, 1900. Originally it consisted of a federation of three socialist bodies, the Social Democratic Federation, the Independent Labour Party and the Fabian Society ; and some trade unions, though the two largest unions, the Miners' Federation and the United Textile Factory Workers' Association, were not represented. An Executive Committee was elected consisting of seven trade-unionists and five socialists.

[1]I should like to thank Mr. Paul Coles, Professor S. E. Finer and Mr. Henry Pelling for their advice about this article.

[2]There is a brief but fairly accurate account in E. Halevy, *Imperialism and the Rise of Labour*, 1929, p. 278. As Halevy suggests, contemporary neglect of the Clitheroe result was almost certainly due to the King's illness and the Coronation.

[3]See J. Keir Hardie, "The Independent Labour Party," *Nineteenth Century*, April, 1903, p. 686.

James Ramsay MacDonald, a prominent member of the I.L.P., was chosen as secretary.

The L.R.C. was intended, at least by some of the socialists, to be a departure from previous organisations, such as the Labour Representation League and the Labour Electoral Association, which had sponsored the election of trade-unionists to the House of Commons. This intention was made clear in the inaugural conference's resolution, moved by Keir Hardie, the virtual leader of the I.L.P., "in favour of establishing a distinct Labour group in Parliament who shall have their own Whips. . . ." Such a decision presaged a degree of independence for the L.R.C., which had not hitherto been attained. It is doubtful whether all of the trade-unionists understood the importance of this point. To many of them, as to most people, the L.R.C. was merely the successor of earlier Labour electoral bodies. Moreover readiness "to co-operate with any party which . . . may be engaged in promoting legislation in the direct interests of Labour," as the resolution ran, must inevitably have conveyed to them a continuance of the policy of working with the Liberal party which had been the practice of the "Lib-Labs." And this impression cannot have been dispelled after the General Election of October, 1900, when a "Labour group" of two was returned to Westminster. For though Keir Hardie, who was elected at Merthyr Tydfil, was a champion of "independence," Richard Bell, the secretary of the railwaymen, who had stood on the same platform as a Liberal candidate in the two-membered constituency of Derby, was more susceptible to Liberal pressure and more amenable to Liberal advances. Furthermore, Bell, as a trade-union secretary, was much more concerned with securing legislation which would re-affirm the hitherto accepted position of the unions, than with maintaining the independence of the "Labour group."

The recurring problem of the L.R.C. was its relationship with the Liberal Party, whose unity was endangered by the stresses of the Boer War, which had broken out in October, 1899. The anti-Imperialism of the Socialists, which was largely shared by trade-union leaders, threw them into a natural

association with the "pro-Boers," the radical wing of the Liberals led by Lloyd George. This alliance, cemented by the subventions of George Cadbury, the prominent Quaker of Bournville, to the funds of both parties,[1] was given practical expression by the I.L.P's "White list" of 1900.[2] Such an understanding between Labour and Liberalism received the commendation of Ramsay MacDonald,[3] while on the other side, Herbert Gladstone, chief Liberal Whip, with the approval of Campbell-Bannerman, the Liberal leader, strove his utmost to come to terms with the L.R.C.[4] As these three knew the chief obstacle to a national agreement was the attitude of the local Liberal associations to working-class candidates. The leaders of the L.R.C., almost without exception, had suffered rejection by the Liberal "hundreds." This had been an important factor in the rise of independent labour representation.[5] But the Liberal caucuses for their part persisted in their opposition to trade-union nominees after 1900, in spite of the threats and blandishments of party headquarters.[6] Consequently between the local Liberal parties and the L.R.C. there was bound to be friction, for it was not at Westminster, but in the constituencies, that the idea of independent labour representation was making its greatest progress. There, MacDonald was indefatigably striving to convince the trade unions of the advantages of joining with local socialist societies to form local Labour representation committees[7] and there the Liberal Chief Whip was endeavouring to persuade Liberal caucuses to accept Labour candidates. Therefore the political interest of the years before 1906 lies especially in the industrial constituencies, where a series of by-elections, of which Clitheroe was perhaps

[1]See Herbert Gladstone Papers (B.M. Add. MSS.) 46046 and *I.L.P. News*, Oct., 1900.

[2]*I.L.P. News*, Aug., 1900.

[3]*Ethical World*, 13th Oct., 1900.

[4]See correspondence about the N.E. Lanarkshire by-election of Sept., 1901, in both the Herbert Gladstone and Campbell-Bannerman Papers (B.M. Add. MSS.).

[5]Henry Pelling, *The Origins of the Labour Party*, 1954, p. 236, ff.

[6]See Herbert Gladstone Papers 46049 and Campbell-Bannerman Papers 41216.

[7]See P. Snowden, *Autobiography*, 1934, Vol. I, p. 103.

the most important, epitomised the difficulties of Liberal-Labour relations.

The Clitheroe division had been carved out of the old two-membered N.E. Lancashire constituency by the Redistribution Act of 1885. It consisted of a string of industrial towns and villages surrounding the wild moorland area of Pendle, Ainsworth's "Lancashire witch country." Of its 19,400 voters, 2,356 lived in Clitheroe, a small industrial town to the North of Pendle, while Padiham an even smaller industrial town to the West claimed 1,400. The valley running South-West to North-East from the direction of Burnley, contained the greater part of the constituency's voters, with 6,797 in Nelson, 2,947 in Colne and 1,560 and 1,224 respectively in the industrial villages of Brierfield and Barrowford. The staple manufacture of all these towns and villages was cotton goods. There were few other large industries besides weaving though some coal was mined.

Since its creation the constituency had been represented at Westminster by a Liberal, Sir Ughtred Kay-Shuttleworth,[1] the son of Sir James Kay-Shuttleworth, the educational reformer. Sir Ughtred had not been opposed since the General Election of 1892, when he defeated the Conservative candidate, W. E. Briggs. Thus the Clitheroe division was an outpost of radicalism in a Lancashire which was overwhelmingly Conservative.[2] This radicalism may be separated into three strands in reality inextricably entwined; the vitality of its non-conformity, the vigour of its socialism and the militancy of its trade-union movement.

The militancy of the constituency's trade-union movement can only be explained in terms of the economic and industrial situation on both a regional and a local level. The division's eighteen thousand textile operatives belonged to one of the

[1]Entered Parliament for Hastings in 1869. 1886 Parliamentary Under-Secretary for India and then Chancellor of the Duchy of Lancaster. 1892–5 Secretary to the Admiralty.

[2]Out of thirty-five seats in what may be called East Lancashire (including Manchester), twenty-seven were Conservative.

"old" unions, the United Textile Factory Workers' Association, which under the leadership of James Mawdsley,[1] the secretary of the spinners, had acquired a reputation for sobriety, moderation and opposition to industrial unionism and its concomitant socialism.

Founded in 1886 in order to promote legislation, the U.T.F.W.A. consisted of a federation of three cotton unions; the Amalgamated Association of Operative Cotton Spinners, the Amalgamated Association of Card and Blowing Room Operatives and the Northern Counties Weavers' Amalgamation. Except for parliamentary legislation, the constituent unions were in charge of their own affairs and they in turn left a great deal of autonomy to the local branches, who were able to affiliate themselves to outside organisations without the permission of the central body. Thus though the loose federal structure was detrimental to the formation of a national or regional policy, it did allow the local unions to be permeated gradually by new ideas and to be infiltrated by their purveyors.

The U.T.F.W.A. had reacted to the proposals for the L.R.C. without enthusiasm.[2] Political divisions within its

[1]It might appear that the death of Mawdsley in Feb., 1902, removed one of the greatest obstacles to the adoption by the cotton workers of a policy of parliamentary representation. Mawdsley who had been a member of the T.U.C. Parliamentary Committee from 1882 to 1897, was certainly no friend of the I.L.P. As his colleague Thomas Ashton said: "He was above all things a business man . . . A lot of silly trash, he said wanted knocking out of the heads of some trade-unionists and as long as he had the opportunity he was determined to do it." (*Cotton Factory Times*, 7th Feb., 1902). However, these views did not preclude a support for labour representation if such a course would profit the textile operatives. Thus Mawdsley had stood as a Conservative at Oldham in a by-election in 1899. His opposition to labour representation was based on a belief in its impracticability and financial wastefulness. Though he did not greet the birth of the L.R.C. with any enthusiasm (see *Sunday Chronicle*, 4th March, 1900), his real opposition was not to labour representation as such, but to the Liberalism with which most working-class M.P.s were allied. He belonged to that school of Lancashire working-men which had adhered to the Conservative party from a detestation of capitalistic liberalism. Therefore it is not surprising that such socialist sympathies as he had, were directed towards the Social Democratic Federation rather than the Independent Labour Party. One of Mawdsley's last acts, was to send a message of support to Harry Quelch, S.D.F. candidate at a by-election at Dewsbury. (See *Justice*, 8th Feb., 1902).

[2]*Cotton Factory Times*, 1st Sept., 1899.

ranks.[1] had prevented it from undertaking trade-union representation as had the miners in Durham. Moreover, as a "pressure group," the union had been largely successful in furthering the interests of its members.[2] However in 1899 was passed the "half-timers'" Bill, which raised the age limit at which children could go to the mill from eleven to twelve. This reform, which deprived some weavers' families of part of their earning potential at a time of increasing prosperity, was strongly opposed by the U.T.F.W.A. and especially by the weavers,[3] both at T.U.C. conferences and at Westminster. The failure of lobbying to avert the passage of this moderately progressive measure, undoubtedly strengthened the case for parliamentary labour representation.[4] Another contributory factor was the fear that employers were federating for political action [5] and combining in trusts.[6] An attack on the legal status

[1]In 1895 parliamentary representation had been discussed by the U.T.F.W.A. Two candidates were chosen, James Mawdsley of the spinners who was to stand as a Conservative and David Holmes of the weavers, who was to be a Liberal. However disagreement over their salary amongst other things caused the scheme to be dropped. (See Webb, Collection D. 275 and 7 B. xcv. (3) and *Cotton Factory Times*, 14th Nov., 1902). Mawdsley stood as a Conservative alongside Sir Winston Churchill in a by-election at Oldham in July, 1899. He was bottom of the poll, but with almost the same votes as Sir Winston, which suggests that his position as a trade-union official can hardly have affected the voting. (See *Oldham Daily Standard*, July, 1899). The resolution on labour representation at the 1899 T.U.C. was virtually opposed by T. Ashton, President of the spinners, when he said : "If our society was to move in politics it would go down immediately." (*T.U.C. Annual Report*, 1899, p. 65).

[2]See S. & B. Webb, *Industrial Democracy*, 1897, p. 260.

[3]There was a higher proportion of child labour on the weaving side of the industry. The Nelson weavers were not so overwhelmingly opposed, because in Nelson the proportion of child labour was smaller. (See Supplement to the Annual Report of the Chief Inspector of Factories and Workshops, 1900, Cd. 841).

[4]*Cotton Factory Times*, 6th Jan., 1899.

[5]The Cotton Employers' Parliamentary Committee had been founded in 1899 to deal with industrial legislation.

[6]For a general discussion of this factor see Henry Pelling, "The American Economy and the Foundation of the British Labour Party," *Economic History Review*, Aug., 1955, p. 1. For more specific reference to the cotton industry and especially to its recession in the early months of 1902 see *Economist*, 12th April, 1902. Also *Textile Manufacturer*, 3rd May, 1902, quoting *Textile World* of Boston, Mass. on the British cotton industry : "Both mill owners and operatives are alarmed at what they feel is the beginning of a decline that promises to end in placing England in a subordinate rank in cotton manufacturing." Finally *Cotton Factory Times*, 2nd May, 1902 : "Cotton syndicates in New York . . . are giving us a lesson, to combat which our workmen are on short time . . . But we are all going the same way."

of trade unions completed the picture of a concerted employer offensive, the brunt of which was borne by the weavers, who in September, 1901 were restrained by an injunction from picketing during a strike at Blackburn.[1] The Northern Counties Weavers' Amalgamation was faced with a bill of about £11,000, while the strikers themselves must have lost £5,000 to £6,000 in wages.[2] Moreover, it appeared that the employers had through the courts the right of access to all papers lying in the society's offices.[3] This situation obviously could only be changed by legislation. But in the circumstances the passage of a remedial statute was impossible, so that it was no wonder that meeting at Blackburn on 11th January, 1902, a conference of the N.C.W.A. instructed its Central Committee "to take the necessary steps to secure direct representation in Parliament." [4]

This decision was a triumph for the Colne and Nelson Weavers whose policy had long been at variance with the unrepentant Gladstonian Liberalism of the union [5] and of

[1] An account of the proceedings can be found in the *Manchester Guardian*, 18th and 25th Sept., 1901 and the *Labour Gazette*, Oct., 1901, p. 295. The legal significance of this and other similar judgments is referred to in the *Report of the Royal Commission on Trade Disputes and Trade Combinations*, Cd. 2825, 1906.

[2] *The Times*, 4th March, 1902.

[3] *Blackburn Weekly Standard and Express*, 15th March, 1902.

[4] *Manchester Guardian*, 14th Jan., 1902.

[5] Probably the Liberalism of the weavers was due to an identification of their interests with those of their employers. The weaving towns were usually smaller and farther from Manchester than the spinning towns. The weaving firms were smaller than the spinning firms. Thus on the weaving side of the industry, personal relations between employer and employees were closer. (See the account in 1886 of Bacup in B. Webb, *My Apprenticeship*, 1926, p. 143 : "Trade unionism is not strong here, class spirit hardly exists because there is no capitalist class . . . ") However the later nineteenth century trend was probably spreading to the weaving towns. As Bruce Glasier of the I.L.P. said at Colne in Sept., 1902 : "Many in that district will still remember some of the old stamp of manufacturers . . . men who had been working men or whose fathers had been working men . . . They spoke in the Lancashire dialect and drank beer in the same public-houses as their workmen. The sons of manufacturers in recent days, however, had been sent . . . to the universities and they had lost the Lancashire dialect. There had been a remarkable change." (See *Nelson Leader*, 19th Sept., 1902).

David Holmes, its President.[1] The militancy of these two branches [2] had been displayed in their support for the Eight Hour Day and labour representation.[3] The delegate of the Nelson weavers had at the T.U.C. in 1899 supported the resolution calling for a labour representation conference, to which in February, 1900, alone among cotton unions, both branches had sent representatives. The Colne Weavers affiliated to the L.R.C. in January, 1901.[4]

Labour representation as an electoral principle was no novelty in the area. In Nelson it dated from the granting of a charter in 1890 and the formation of a Trades Council. However, due to the opposition of the socialists, weavers'

[1]W. Bennett, *The History of Burnley*. Burnley, 1951, Vol. IV, p. 128. Holmes was no friend of the L.R.C. When interviewed about the proposal for its inauguration he said : " . . . the Socialists are trying to get hold of the funds of trade unions and co-operators to use them for their own purpose." (See *Nelson Chronical*, 15th Sept., 1899).

[2]There are many possible reasons for this distinctiveness. The proportion of females to males employed in cotton was much lower in Nelson and Colne than in other textile towns. (See S. J. Chapman, *The Lancashire Cotton Industry*, Manchester, 1904, p. 60). In Nelson the trade unions had established an ascendancy due to the very recent expansion of the town, whose population had grown from 3,500 in 1865 to 32,000 in 1902. Thus in very recent times the area had been inundated by an influx of population from the textile districts of Yorkshire, usually young married people who were looking for better conditions. (See *Nelson Gazette*, 1st Jan., 1952. "The Rise of Labour" by Ald. Smith, C.B.E., J.P., Chairman of the Lancashire County Council : "They were pioneers, emigrants, seeking something better than the conditions from which they had fled—and more receptive, consequently, to ideas which were new and idealistic than was common with working people in those days.") Perhaps the situation of these migrants, more akin to that of the city proletariat, who could have few personal contacts with their employers, contributed to the bitterness with which the industrial struggle was waged in Nelson in the 1880's and 1890's. (See *The Origin and Progress of the Nelson Weavers' Association*, 1870–1920. Burnley, 1922). It is difficult to say how far Nelson affected the position in Colne, a much older town with Chartist memories.

[3]The Nelson Weavers first sent a resolution on the Eight Hour Day to the U.T.F.W.A. in Feb., 1895. (See *The Origin and Progress of the Nelson Weavers' Association*, p. 21). They first raised labour parliamentary representation in 1894 (*op. cit.*, p. 20). See also Minutes of a Members' Meeting of the Nelson Weavers' Association, 30th Aug., 1899, when it was resolved that the union should support labour representation and the abolition of capital punishment and oppose the raising to fourteen, of the age at which children could go to the mill.

[4]*Cotton Factory Times*, 11th Jan., 1901.

representation on the Town Council was not a great success.[1] On the other hand in Colne, with the leadership of their able secretary, A. B. Newall, the weavers were much more successful in obtaining municipal representation. In 1898 Newall was elected for the Central Ward and in succeeding years candidates variously described as "Weavers'," "Labour," "S.D.F." and "Socialist" were returned to the Town Council.

It was not until 23rd September, 1901 that a general meeting of the Nelson Weavers' Association decided to depart from its policy of non-participation in elections and to run candidates at the municipal elections in the coming November. Two of their candidates were returned—a success which gave rise to increased aggressiveness, one self-declared supporter of labour declaring: "... We fully intend to go against Liberalism on all occasions." [2] The Nelson trade union movement was thus at last taking up the position of labour at Colne, but it was left to the Trades Council of the latter town at a meeting on 10th December, 1901 to raise the question of a local conference on labour representation. The matter was referred to affiliated societies,[3] but at a further meeting of the Trades Council on 7th January, 1902 it was agreed "to join hands with the Nelson, Colne, and District Textile Federation in arranging a conference and public meeting." [4]

The division owed its flavour of sturdy dissent above all to the strength of the Independent Methodist Connexion founded in 1808,[5] though the Colne and Nelson Churches, break-aways from the Primitive Methodists and originally known as the "Free Gospellers," had not become members of

[1]On 13th Jan., 1896, the annual meeting of the Nelson Weavers decided by 55 votes to 25 not to take part in any future parliamentary or local election. (See *Nelson and Colne Express*, 18th Jan., 1896). When the question of support for their secretary as a municipal candidate, was raised at the quarterly meeting on 16th Oct., 1899, the opposition of the Socialists led to his withdrawal from the contest. (See *Nelson Chronicle*, 20th Oct., 1899).

[2]*Nelson Chronicle*, 8th Nov., 1901.

[3]*Cotton Factory Times*, 13th Dec., 1901.

[4]*Ibid.*, 10th Jan., 1902.

[5]The five societies at the inaugural conference in 1808 were: Macclesfield, Manchester, Warrington, Stockport and Oldham. They represented 1,219 members. (See J. Vickers, *History of Independent Methodism*, Bolton, 1920, p. 20).

the Connexion until 1855.[1] By 1902 there were 1,329 members and 17 chapels in the two towns making the circuit the strongest in a denomination which had only about 8,500 members throughout the whole country.[2] While not entirely pacifist in sentiment the Independent Methodists were definitely anti-militarist,[3] their opposition to the Boer War increasing with its prolongation.[4] However, the significance of this factor can be over-estimated. There was certainly no overt demonstration against the war in North-East Lancashire,[5] while non-conformity in the country as a whole was very much divided,[6] On the other hand, non-conformity was united in its attitude to liquor, an important political issue at this time,[7] and there can have been few localities in which the feeling was as strong as at Nelson and Colne. Here the Independent Methodists, provoked by the opening of more drinking clubs locally,[8] were in the forefront of a campaign against the sale of alcohol.[9]

The ideological rift on the issue of "the drink" had undoubtedly shaped the struggle in the Clitheroe division,

[1] *op. cit.*, p. 167.

[2] *Colne and Nelson Times*, 20th June, 1902.

[3] Some members of the Society of Friends in the Warrington area threw in their lot with the Independent Methodists. (See W. J. Townsend, H. B. Workman and G. Eayrs, *A New History of Methodism*, 1901, Vol. I, p. 559).

[4] See *Independent Methodist*, Jan., 1900, in which the editor, W. Brimelow, though not exonerating the Boers, maintained that Britain was seriously to blame for the war. In the following month he alleged that the British nation was "blinded by the pride and arrogance of Empire." The annual conferences of the sect in 1900 and 1901 unanimously adopted resolutions deprecating the growth of militarism and calling for the settlement of international disputes by arbitration. (*Ibid.*, July, 1900 and July, 1901).

[5] However, see *Nelson Chronicle*, 4th Oct., 1901, where a correspondent complained bitterly of "the political parson in Nelson," who disseminated "pro-Boer" sentiments from the pulpit.

[6] M. Edwards, *Methodism and England*, 1943, p. 161.

[7] J. Rowntree, *The Temperance Problem and Social Reform* had been published in April, 1899 and by Dec., 1900 had run to nine editions. See his ingenious table, pp. 681–687 in which he computed that assuming that at the General Election of 1895 each "on" licensed house caused the transfer of two votes from the Liberals to the Conservatives, 152 seats would have been lost to the latter party. In Lancashire the Conservatives would have won 22 seats less.

[8] See *Nelson Chronicle*, 4th Jan., 1901, which reported that the number of drinking clubs in Nelson had increased to forty-four over the previous three years.

[9] *Independent Methodist*, 1900 and 1902.

between the contending socialisms of the Social Democratic Federation and the Independent Labour Party. The former body, secular and revolutionary, had been founded in London in 1881. The latter, more pragmatic and evolutionary in its approach, dated from the Bradford conference of 1893. While the secularism of the S.D.F. made its converts in London and Lancashire,[1] the I.L.P. appealed to the dissenting areas of the provinces. Of all industrial areas, Lancashire, with its links between the Conservative party and the working-class, forged by the agitation against the Liberal Poor Law and fashioned from Conservative Working Men's Clubs in which strong liquor was consumed ; was the one in which the I.L.P. had made least progress. In Yorkshire, on the other hand, a stronghold of nonconformity, it was natural that the I.L.P. should be the stronger socialist party.[2] It is not surprising therefore that the clash between the two bodies should take place in the nonconformist areas of Lancashire. The fact that the Clitheroe division was on the border between the two counties emphasised the transition, which was geographical as well as ideological.

Yet up to the end of the century the two socialist parties had succeeded in working together amicably in the district.[3] They ran a joint paper, the *Socialist Journal* and they co-operated at School Board Elections. However, even before the national quarrel took place, the breach between these two organizations was implicit in their differing methods and attitudes. The S.D.F. stressed agitation and propaganda. The I.L.P. relied on education.[4] Moreover, while S.D.F. clubs

[1]H. W. Lee and E. Archbold, *Social Democracy in Britain*, 1935, pp. 121–122.

[2]For example, while the Bradford branch of the I.L.P. had 340 members in 1902, Blackburn had only 82. Halifax claimed 240 members as against Manchester's 96. Nelson I.L.P. at this time had 72 members. (See *I.L.P. Conference Report* 1902).

[3]Discussion for a fusion between the I.L.P. and the S.D.F. had taken place in 1897 and continued until 1902. (See *I.L.P. Annual Conference Report* 1899 and *S.D.F. Conference Report* 1902). At these conferences the Lancashire members of both parties, such as Higham of the Blackburn I.L.P. and Irving of the Burnley S.D.F., supported fusion.

[4]Ald. Smith, *loc. cit.*, 8th Jan., 1952 refers to the lectures in Nelson of Katherine St. John Conway. a Newnham graduate, and those of her future husband, Bruce Glasier. He considers it was their "deep humanitarian note" which led him to join the I.L.P. rather than the S.D.F. The latter "tended to antagonise . . . people who were sympathetic to their ideals."

sold intoxicating liquor,[1] I.L.P. branches remained true to "temperance." [2] Ultimately it was this factor which more than anything else swayed this predominantly dissenting constituency to the I.L.P. type of Socialism.

It is, therefore, no coincidence that Philip Snowden, a West Riding man who had lived in Nelson and was the I.L.P's foremost temperance advocate, should have been the real architect of the political Labour movement in that part of North-East Lancashire.[3] Snowden was the son of a Cowling [4] weaver who had moved to Nelson in 1879, in search of work like so many other Yorkshire families at that time.[5] Although he returned to Cowling in 1889 when his father died, Snowden retained his links with Nelson and Colne and in these towns became the rousing evangel of the I.L.P. with his promise of "the Christ that is to be." In every way he was marked out to be the first Labour candidate for the division.

The I.L.P. Council had resolved to recommend Snowden for Clitheroe if he were not successful at Blackburn.[6] However the first local conference of labour and socialist organisations on 31st January, 1902,[7] meeting amid the usual obstructionist

[1] *Nelson Leader*, 17th March, 1902. "Barrowford Jottings": "Somehow secularism seems to cling to socialism like ivy to the oak. The attitude of many of the socialists to temperance reform has provoked an attitude far from friendly among temperance reformers. We have an idea that many of the socialists recruited from the Tory ranks retain some of the old Tory leaven." Snowden was specifically exempted from these strictures.

[2] See *I.L.P. News*, Oct., 1899. A report of the Nelson branch, containing a glowing account of its activities, concluded : "The club does not sell alcohol and its members are on excellent terms with each other and their fellow citizens."

[3] See his *Socialism and the Drink Question*, 1908.

[4] Five miles North-East of Colne.

[5] Snowden's father had been a Chartist. See P. Snowden, *op. cit.*, Vol. I., p. 19.

[6] Minutes of the I.L.P. Council at Glasgow, 16th Sept., 1901.

[7] *Cotton Factory Times*, 7th Feb., 1902.

tactics of the S.D.F.,[1] adjourned until after the L.R.C. annual conference at Birmingham. When it did assemble the local conference conformed to pattern. At a public meeting at Nelson on the evening of 28th February, it was moved and carried :— "That in the opinion of this meeting the time has arrived when Labour should be directly represented in Parliament, and in all public bodies, on independent lines." Councillor Hodge of the L.R.C., in support of the resolution, spoke of his frustrating experiences as a Liberal, referred to the growing danger from trusts and hoped his audience would throw off "the shackles of the partyism of political parties . . . and rally round the banner of labour." [2] At the same time at Colne, Keir Hardie addressed a similar audience urging them to use the "great facilities at their command." [3]

On the following day the conference met at Nelson under the Chairmanship of A. B. Newall of the Colne Weavers.[4] Trade Unions, Co-operative societies and branches of the S.D.F. and the I.L.P. in the division were represented among about a hundred delegates.[5] Four resolutions were discussed. The first, which was unanimously adopted, welcomed the formation of the L.R.C. and pledged the conference to advance its interests especially among trade-unionists. The second referred to the conditions under which a Parliamentary candidate should stand. The I.L.P. moved an amendment for "a Labour and Socialist candidate," a formula which had been used by Snowden at Blackburn in 1900 and which was intended to placate the S.D.F. When their amendment was defeated the

[1]The national S.D.F. had decided in Aug., 1901, to disaffiliate from the L.R.C. on the grounds that it was not sufficiently class-conscious. In Nov., 1901, the S.D.F. had queered the pitch of the I.L.P. at the Dewsbury by-election. In the Clitheroe division the S.D.F. in Jan., 1902, decided to bring out its own candidate and invited the constituency's trade societies as delegates to a conference at Nelson on 25th Jan., 1902. For obscure reasons, so characteristic of the S.D.F., this conference was never held.

[2]*Nelson Leader*, 7th March, 1902.

[3]*Cotton Factory Times*, 7th March, 1902.

[4]He had represented them at the 1902 Conference of the L.R.C.

[5]The basis of representation was two delegates up to 500 members ; each extra 500 members one delegate more ; but no society to send more than six delegates. (Infancy of Labour, Vol. I, f. 96. Brit. Lib. Pol. Sci. The two volumes with this title comprise the minutes of the L.R.C. and other relevant documents between 1900 and 1912).

I.L.P. did not withdraw.[1] They accepted the wish of the majority thus aligning themselves with the trade-unionists and against their erstwhile socialist allies. The S.D.F's delegation moved most of the amendments, characteristically explaining however, that they were there with a watching brief and could not promise to be bound by the decisions arrived at, even if their amendments were carried. Ultimately the resolution adopted stipulated that candidates should : "Loyally co-operate with a Labour party in Parliament, in advancing the interests of Labour, and that on all Labour matters they should act together, independently of all other parties." The third and fourth resolutions which were passed after a little discussion, urged the delegates to encourage their societies to affiliate to the L.R.C. ; and recommended the various societies to consider ways and means of raising funds for election expenses and for maintaining an M.P. For the purpose of giving effect to the last resolution the Conference was adjourned for three months.[2] Thus no candidate was adopted and this may well have been because Philip Snowden, who was everywhere spoken of as the likely choice, was at that moment fighting a by-election at Wakefield. A sequel to this conference was that a sub-committee was set up of the Colne Trades Council and the Nelson and District Textile Trades Federation, with the respective secretaries of the two bodies, A. B. Newall and W. Ward, as joint secretaries. Meeting on 14th March, 1902 this sub-committee discussed the conference which had been held, and fixed the date of its successor as 28th June, 1902.[3]

The period between the two conference was one in which the trade societies took their decisions. Some indication of the way in which opinion was moving may be gathered from the first quarterly report of the Nelson Weavers for 1902 which ended : "Therefore, let us as workers sink our little differences and go hand in hand and return representatives to Parliament and on all public bodies from our own class, and show the capitalist class that we are determined not to have them as our representatives any longer." [4] At a members' meeting of the

[1] *Colne and Nelson Times*, 25th July, 1902.
[2] *Labour Leader*, 8th March, 1902.
[3] *Cotton Factory Times*, 21st March, 1902.
[4] *Nelson Leader*, 11th April, 1902.

Nelson Weavers on 12th May, 1902 a resolution was passed instructing the Association to take a ballot on the question of a levy for the purpose of securing a labour candidate. When it was held the figures confirmed the attitude of the leadership. The Colne and Clitheroe Weavers had already decided by large majorities to take this step.[1] The sub-committee now issued the agenda for the conference, at Colne on 5th July, 1902 which consisted of the two questions : Shall levies be on all members or only on male members? What amount per member? and a recommendation that in the event of the formation of a Clitheroe Divisional Labour Party, a sub-committee should be appointed for the conference to draw up a code of rules.[2]

The rise of labour as a political force had not been entirely neglected by local Liberals. The *Nelson Leader*, mouthpiece of Liberal Radicalism in the constituency, was in favour of some arrangement with the "Labour party," though most of its efforts on behalf of such an agreement were devoted to persuading working-men that their real interests lay with the Liberals. During 1901, the journal's predecessor, the *Nelson Chronicle*, had published a series of articles on "Liberalism in the Clitheroe Division," which deplored the moribund state of the Liberal organisation and drew unfavourable comparisons between Liberal propaganda and the programme of education being carried out in the division by the I.L.P., the Ruskin Hall movement and the Co-operative societies. The anti-Liberal bias of the latter was exemplified by a lecture in which Councillor James Johnston of the Manchester I.L.P. went so far as to illustrate the attitude of the Liberal party to social reform, with a lantern slide of Sir William Harcourt drinking champagne at a public dinner, while a half-clothed little girl sold matches in the street outside.[3]

Of course, disagreements over South Africa at a national level were inevitably reflected in the constituency.[4] Sir Ughtred Kay-Shuttleworth took his position behind that of

[1] *Ibid.*, 30th May, 1902.
[2] *Ibid.*
[3] *Nelson Chronicle*, 15th March, 1901.
[4] *Ibid.*, 25th Oct., 1901. "Party discipline has become relaxed over the question of the South African War."

the leader of the party, Campbell-Bannerman.[1] He was deeply shaken when in February, 1902, Lord Rosebery vowed his "definite separation" from his leader[2] and promptly dissociated himself from that epistle,[3] though Alderman Catlow, the former Mayor of Colne and President of the Clitheroe Division Liberal Association, objected to committing himself, beyond admitting that "the Earl lacked backbone."[4] However dissatisfaction both with the local organisation and with the member was still widespread. Criticism of the caucus appeared in the local press[5] and though the local association defended its M.P.,[6] that did not prevent attacks being made on him also. Complacency, apathy and neglect of the local Liberal agent[7] were the chief accusations made by a rank and file that had become apprehensive of the Labour party.[8]

These forebodings seem to have been forgotten when, on 26th June, 1902, the morning papers announced that Sir Ughtred had been granted a peerage in the Coronation Honours List. The Liberal *Colne and Nelson Times* asserted: "The Liberal Party is . . . in a perfectly equipped state, ready for the writ, however soon it may come. The supreme issue on which the election will be fought by the Liberal Party is the bread tax to which the Government have committed themselves."[9]

However, the necessity of conciliating labour, as his correspondence with the Liberal Chief Whip reveals, was not

[1] *Ibid.*, 8th Nov., 1901. An account of a "political high tea" at which Sir Ughtred praised their leader. The meeting carried a vote of confidence in its M.P.

[2] See Rosebery's letter to *The Times*, 21st Feb., 1902.

[3] *Nelson Leader*, 28th Feb., 1902.

[4] *Ibid.*, 14th March, 1902.

[5] *Ibid.*, 31st Jan., 1902. "A more inert body of leaders I have never met."

[6] *Ibid.*, 27th March, 1902.

[7] *Nelson Chronicle*, 13th Dec., 1901.

[8] See *Nelson Leader*, 14th Feb., 1902. For an account of a Liberal meeting at which one member said that if the trade-unionists brought out Snowden, the Liberals would have to fight hard for the seat.

[9] The narrative of the next few weeks is based, unless otherwise stated, mainly on the *Colne and Nelson Times* and the *Nelson Leader* and to a less extent on the *Burnley Express*, the *Manchester Guardian*, the *Daily Dispatch*, the *Cotton Factory Times*, the *Labour Leader* and the Herbert Gladstone Papers 46050.

far from the retiring member's mind.[1] Though Alderman Catlow did not appear to feel this need so strongly,[2] his appreciation of the difficulties was demonstrated by his journey to London to discuss possible candidatures with Herbert Gladstone. Following this conversation, at a meeting of the local Liberal Executive on 27th June, the names of Augustine Birrell, the Hon. Philip Stanhope and Mr. Franklin Thomasson of Bolton, were sent to an interviewing sub-committee. The Liberal Executive at this time had no cause to doubt the reality of a Labour candidature. It had been reported that the local I.L.P. had recommended Snowden to the L.R.C., while the secretary of that body had written to the press emphasising their determination to contest the seat.[3]

Meanwhile Birrell had withdrawn and on 1st July, the Liberal sub-committee asked Stanhope to appear before them on the following day. Excitement mounted when he arrived at Colne next morning. Though he was the youngest son of the fifth Earl of Stanhope and brother of Edward Stanhope, a former Conservative Minister of War, Stanhope had earned for himself the reputation of being one of the most advanced of all British Radicals. Perhaps he had inherited some of the progressive ideas of his grandfather, "Citizen Stanhope," who had taken a leading part in obtaining measures of relief in favour of non-conformists. Entering the House of Commons in 1886 for Wednesbury when he was thirty-nine, he had been one of the most active critics of the Factories Act into which he had endeavoured to insert amendments in connection with the nail and iron trades of the Black Country. He was greatly interested in arbitration as a solution of international problems. As President of the National Reform League he participated in a national campaign against the war in South Africa ; while in Parliament, until he lost his seat in 1900, he was one of the leaders of the Radical attack on the Government. His anti-militarism commended him to the I.L.P., his advocacy of factory reforms earned him the gratitude of the trade-unionists,

[1]Campbell-Bannerman Papers 41221.

[2]Catlow favoured J. A. Bright, the son of the illustrious John Bright. Sir Ughtred referred to trade-union questions : "on which the Bright firm used to be rather unbending."

[3]*Daily Chronicle*, 27th June, 1902.

so that his probable choice by the Liberals was a move more calculated than any other to conciliate the labour and socialist forces in the constituency. On arrival Stanhope willingly expressed his views on the relations between Liberalism and Labour. Declaring himself a supporter of a Liberal-Labour alliance, he suggested that the reduction of election expenses and the payment of M.P.s was the first step to be taken to facilitate the representation of labour. He believed that immediate action should be taken to reverse the effect of the Taff Vale judgment.

On the afternoon of 2nd July, local labour leaders met the Liberal sub-committee in an attempt to find an amicable solution. Alderman Catlow who presided, in his opening remarks referred to the prospective improvement in the Liberal party and pointed out what a pity it would be if the division between the Progressive forces in the constituency let in a Conservative. One of the labour leaders asked if any nomination of the Labour party would be entertained, or did the Liberal committee claim the right to nominate. Alderman Catlow replied that they did not entirely claim the right, but they thought they had a prior claim as they were in possession of the constituency. The Liberals had been strongly influenced by the position of the Labour party and they were taking care that no one should be considered, who would not advocate the claims of the Labour party on all questions that would come before Parliament. They felt satisfied any name they had chosen could do all that a labour candidate could do, but they certainly could not profess to object to any name the labour men might advance. Alderman Catlow asked the labour leaders whether, if the Liberals considered any name they cared to mention, would he if elected, after being allowed full powers to act freely on labour questions, be answerable to the Liberal whip? To this the labour leaders gave a decided negative and refused to commit themselves to any definite course of action before the conference on 5th July. The meeting then broke up. A. B. Newall, who was leader of the labour delegation said after the meeting: "Well, if the Liberal party are prepared to do what their Chief Whip has said so often, to earmark constituencies for a Labour candidate,

they have now a grand opportunity to show whether they are sincere or not . . . There are no less than 18,000 trade-unionists in the Clitheroe Division, and therefore, our contention is that we meet the very first requirements of Mr. Gladstone in being a Labour constituency . . . Our strongest places for the Labour vote are Colne, Nelson, Clitheroe and Padiham."

On the evening of the same day the Liberal selection committee met Stanhope and invited him to submit his name for nomination as a candidate unanimously expressing the opinion that his candidature would meet with the general concurrence of the Liberal and Labour parties in the Division. Stanhope, who had returned to Manchester, wrote to Catlow explaining that he had consented to the submission of his name with the object of reconciling Liberalism and Labour and that therefore he could not take part in a contest against the forces of trade-unionism. In a letter written at the same time to Herbert Gladstone he reported that the labour leaders were "hopelessly intractable" and advocated choosing "some Labour men acceptable to both sides." The retiring member, advised arbitration and accepted Catlow's suggestion that he should not immediately apply for the Chiltern Hundreds, but should await events.

However, on the afternoon of 3rd July, steps were taken which rendered Stanhope's suggestion of a Liberal-Labour candidate and Kay-Shuttleworth's attempt at arbitration untenable. The L.R.C. meeting at Colne received a deputation of trade-unionists from districts in the division and ascertained that there was a determination to contest the seat. Edward Pease, Keir Hardie, Richard Bell and MacDonald were entrusted with the task of drawing up a manifesto in favour of the selected candidate, provided that he conformed to the conditions of the L.R.C.[1] "We are satisfied" said MacDonald afterwards, "that funds will be forthcoming both for the contest itself and for the maintenance of a Labour member in the House of Commons."

The meeting of the Liberal Council on the evening of the same day does not appear to have taken these preparations

[1]Infancy of Labour, Vol I, f. 118.

seriously. They listened to Stanhope's letter and then unanimously decided to invite him to be their candidate. A likely interpretation for their action, is that they were unaware whom the labour conference would select on the following Saturday. As one journal remarked: "Should Snowden be the choice of the delegates it is probable . . . that Stanhope will proceed with his candidature. A Socialist is not regarded as a Labour man purely, by some Liberals, and whilst they may be willing to forego any claims they may have for a bona-fide Labour man, they would not feel the same obligation in dealing with a Socialist." In fact, as we know, both the Liberal Executive and Stanhope were playing a waiting game, and on the afternoon of 4th July, the latter sent a telegram to Herbert Gladstone, which read: " . . . No necessity hurried decision—much better let things develop a little . . . I have by no means abandoned possible candidature."

Meanwhile the Parliamentary Committee [1] of the I.L.P. Council met at Leeds on 4th July, with F. W. Jowett [2] as a co-opted member. With Stanhope's declaration in mind their course was easy: "Snowden should withdraw in Shackleton's favour provided the latter's position on the question of independence was satisfactory . . . " [3] Thus in accordance with their declared strategy, a socialist was withdrawing in favour of a trade-unionist who was willing to pledge himself to be independent.

On Saturday afternoon, 5th July, eighty-one delegates representing 18,000 workers, largely drawn from the textile trades but with a sprinkling of gas workers and printers and the I.L.P. and Socialist District Council, met in the Co-operative Hall at Colne in order to select a Labour candidate. A. B. Newall was in the chair, while W. Ward, secretary of the Nelson Weavers, acted as secretary to the conference, which was also attended by Ramsay MacDonald and James Sexton as representatives of the L.R.C. At the outset of the Conference

[1] A committee for parliamentary candidatures consisting of Bruce Glasier, Keir Hardie, MacDonald and Snowden.

[2] A Bradford Labour councillor and one of the founder members of the I.L.P.

[3] Report of the Parliamentary Committee of the I.L.P. Council to the Council's meeting at Sunderland, 28th and 29th July, 1902.

the Chairman read a letter from Snowden contending that : "Mr. Shackleton, as a trade-union official, and one thoroughly acquainted with the staple trade . . . has a far better claim, and would make a more useful representative." After some discussion Shackleton was adopted as the Labour candidate and for his upkeep if returned it was resolved to make a levy of 6d. per member per year. It was also decided to form a General Election Committee.

David Shackleton, the son of a watchmaker, was born in Rossendale at Cloughfold, between Rawtenstall and Bacup, on 8th November, 1863. He had started work as a half-timer. When he was fifteen his family had moved to Accrington, where in 1888 he became President of the Accrington Weavers. His association with that body had been marked by a complete absence of strikes. In May, 1894, he had been appointed secretary of the Darwen Weavers and there he had taken a prominent part in obtaining a 2½% increase in wages. From 1898 onwards as a member of the Council of the U.T.F.W.A. he had been a member of deputations of textile operatives to the Home Secretary. As Vice-President of the Northern Counties Weavers' Amalgamation he was opposed to the 1899 Half-Timers' Act, but his views were not nearly so extreme nor so vehemently expressed as those of his President, David Holmes. Shackleton supported the eventual raising of the age limit, but was opposed to it being done precipitately.

Shackleton, a genial bearded giant of a man, who was accustomed to defend the Half-Timers' system by reference to his own girth and stature, was incapable of bitterness against anyone. His own rapid progress had probably inclined him to moderation. A Justice of the Peace at twenty-nine, at thirty-two he had become a Darwen Town Councillor as a representative of the Trades Council and served as chairman of the Tramway and vice-chairman of the Building Committee. His economic philosophy and his nonconformist upbringing naturally inclined him to the Liberals. As a youth of seventeen he had joined the Party and for several years he had been secretary of the Liberal Association in the North-East Ward of Accrington. However, since moving to Darwen he had withdrawn from political activity believing that a paid official

of a trade union, whose members had so many different affiliations, should have no part to play in national politics. He had devoted himself to trade-union matters opposing the practice of artificial "steaming" in weaving sheds and "driving" in mills. In 1902 he had been invited to accompany a delegation of cotton manufacturers to the United States to investigate the growing competitiveness of the American textile industry, especially the introduction of the new Northrop loom.

Keir Hardie welcomed the choice of the conference because it proved that : "The I.L.P. is in the L.R.C. not for what it can get out of it, but in order that it may consolidate the Labour movement." The *Cotton Factory Times* characterised the decision as : "The historical starting point of a new movement," declaring with all the fervour of the newly converted : "If Labour is ever to be free it can only be done by strong and determined action." The *Nelson Leader* questioned whether this "unwise step" would meet with the approval of "the general body of trade-unionists throughout the Division"and even asserted that they had not been consulted. The *Colne and Nelson Times*, as usual more realistic in its approach, quoted a correspondent who had written : "No one acquainted with the feeling among the workers in the constituency could have looked for any other result."

The Conservatives were late starters in the electoral race. They had not fought the seat since 1892 and consequently they found it difficult to estimate their strength in the division. At the beginning of 1902 the Conservative Associations at Nelson and Colne had appointed agents, probably hoping to profit from the divisions within the Liberal Party. The prospect of a contest was viewed with optimism by the Conservative chairman, Sir John Thursby. He returned from London to hold the first Executive meeting of the rejuvenated Clitheroe Divisional Conservative Registration Association on 2nd July at Burnley. The meeting which was almost unanimously in favour of a contest, considered the names of two possible candidates, Colonel Parker of Browsholme and T. Garnet of Clitheroe, but "in view of the peculiar circumstances the final decision was adjourned for a few days." These "circumstances"

were doubts about the type of Liberal candidate to be nominated, the contingency of a Labour candidate and the possibility of a three-cornered fight. It seems evident that if an Imperial Liberal had been nominated the Conservatives would have allowed the seat to go uncontested,[1] but the adoption by the Liberals of Philip Stanhope was a challenge to their patriotism.[2] The selection of David Shackleton on 5th July, by threatening to split the left-wing vote, vastly increased the Conservatives' chance of victory and sent Sir John Thursby hurrying to London where on 7th July he met Captain Middleton, the chief Conservative agent. However, because of the Liberals' hesitations no decision was taken and the Conservatives settled down to playing a waiting game.

The adopted Liberal candidate appears to have himself been in a state of indecision. At the week-end (5th to 6th July) he paid a visit to the Reform Club at Manchester where he laid his position clearly before leading Lancashire Liberals. It may have been as a result of this meeting that on 6th July, he sent a telegram to Herbert Gladstone, suggesting Sam Woods, secretary of the T.U.C. Parliamentary Committee, as a compromise candidate. On the following day in a letter to Catlow he quite definitely announced his withdrawal in view of the result of the previous Saturday's conference which he took to consist of "the duly accredited representatives of the organised trade unions. Yet in a letter on 9th July, commenting on the same conference he expressed his doubts about its representative nature. Stanhope mentioned that he had discussed the situation with Campbell-Bannerman during the previous week-end, commended the dilatory tactics of Alderman Catlow and in the postscript anticipated that if there were no Liberal opposition the Tories would fight and win.[3]

[1] *The Times*, 1st July, 1902.

[2] See *Burnley Express*, 2nd July, 1902: ". . . it is somewhat significant that the . . . Liberals should favour . . . a gentleman who has assumed such an unpopular attitude at a time when the very existence of the British Empire was threatened."

[3] Herbert Gladstone, Papers 46051. Stanhope's hesitations were almost certainly due to the decision of the local S.D.F. to submit four names of suitable candidates to London headquarters.

The fact that Shackleton, far from being a Socialist, was indeed a Liberal who had declined the invitation of the Darwen Liberals to contest their constituency in 1900, made the position very awkward for the Liberals, who had not yet grasped the significance of independent Labour candidates, and therefore, could not understand why a locally prominent Liberal should assert that he was "a Labour man purely and simply." The growing impatience of the local Liberal caucus expressed itself in a feeling that the Liberal party should nominate a candidate even if Stanhope withdrew. On the evening of 8th July the Liberal Executive met to discuss the situation. The report of the committee's deliberations was awaited by a considerable number of people as it had been rumoured that Stanhope had sent a telegram declining the association's invitation to stand. After the meeting Alderman Catlow denied that he had received any telegram to that effect and reaffirmed that a Liberal Candidate would certainly be nominated. His concern with the complex situation which had arisen is, however, apparent from a telegram which he sent to Campbell-Bannerman asking for an interview with the Liberal leader and Herbert Gladstone. This interview took place at the House of Commons on the afternoon of 10th July,[1] and gave rise to dark rumours about the intentions of the Parliamentary leaders.[2]

Meanwhile the Labour party had moved into action, with the local weavers' organisations as the highly efficient nuclei of the election committees. On 10th July a meeting at Padiham formed the Clitheroe Division L.R.C., and a council was elected with Councillor White of Colne as Chairman, Councillor Newall as Secretary and W. Ward of the Nelson Weavers as Treasurer.[3] Shackleton began to visit the

[1]There appears to be no account available of the meeting, but what happened can be deduced from later events.

[2]See *Colne and Nelson Times*, 11th July, 1902. "It is stated on good authority that, with a view to the prestige of the Liberal party in other constituencies, the leaders in London are using their influence on local leaders of the Liberal party not to offer any opposition to Mr. Shackleton's election."

[3]Colne still had a preponderance in labour counsels.

constituency, where he was soon feeling the pressure of various interest groups.[1]

On 14th July, four days after the meeting with Campbell-Bannerman and Herbert Gladstone, the Liberal sub-committee met a deputation of the Labour party including Shackleton and Newall at the Reform Club in Manchester, in an attempt at compromise between the two forces.[2] The labour leaders were asked if they would be prepared to accept Philip Stanhope as a Liberal-Labour or a Labour-Liberal candidate. To this question they gave an unequivocal negative. Then the Liberals asked if Shackleton would go before the electorate as a Liberal-Labour candidate. Again the answer was "no." The labour leaders also refused an offer of arbitration. Shackleton declined to promise to state that he was a Liberal in his election address, although he assured them that in that document he would make his principles clear. The Labour party were then asked, if in the event of there being any other contest in Lancashire, they would use their influence to prevent labour opposition. The labour representatives replied that they could not give any undertaking, but they promised to use their influence.[3] This conference which finally dashed all Liberal hopes of a compromise made vacillation on Stanhope's part unnecessary. His letter[4] to Catlow dated 7th July, in which he declined the nomination, was made public on the evening of 14th July. If the Liberal party wished to fight they now had to find another candidate.

The negotiations at the Reform Club had a mixed reception. Liberal feeling nationally was pleased, but the Conservatives interpreted the meeting as a refutation of the L.R.C.'s claim to independence. Bruce Glasier, the Chairman of the I.L.P. wrote to Keir Hardie : "It would be a great misfortune were

[1]The temperance group and the women trade-unionists led by Mrs. Pankhurst. The latter passed a resolution that : " . . . any M.P. for the Clitheroe Division . . . should put the enfranchisement of women in the forefront of his political programme." (See *Nelson Leader*, 18th July, 1902).

[2]Bruce Glasier Correspondence. Glasier to Keir Hardie, 17th July, 1902. "Had one of our men been on the committee that would not have happened."

[3]The account of this meeting is Alderman Catlow's.

[4]Catlow had denied that he had received a *telegram*.

Shackleton to prejudice the independent attitude." [1] Snowden, whom he met in Manchester on 17th July was more optimistic.[2] The two men concocted an imaginary interview between Glasier and an imaginary journalist, which was published on the following day. It emphasised that Snowden had only stood aside because Shackleton had pledged his independence and maintained that in the event of Shackleton accepting some arrangement with the Liberals "not only the I.L.P. but the bulk of trade-unionists would repudiate such a proceeding." On 21st July, the Nelson I.L.P. in a manifesto, revealed that one of their delegates had seconded Shackleton's nomination on 5th July and appealed for support to all Socialists.

On the evening of the day that Snowden and Glasier met in Manchester, the Executive of the Clitheroe Divisional Liberal Association assembled at Padiham in full force. "There were determined party men from the towns, and even the far-away villages sent their representatives." After the meeting Alderman Catlow informed the Press that a meeting of the Liberal "600" would be held on the following Monday, and that the Council would be addressed by F. Thomasson of Bolton, the son of a cotton manufacturer, who had been prominent in several movements for social reform. The *Nelson Leader* interpreted this as "tantamount to a decision to fight," adding that the caucus "seems to us to be strongly in favour of fighting."

However, yet more disappointment was in store for the Liberals. On Friday, 18th July a deputation went to see Thomasson who, influenced by the attitude of the Liberal leaders in London, gave no hope of his contesting the seat.[3] His withdrawal no doubt greatly helped the Liberal Executive meeting on 21st July to recommend to the Council that evening that Shackleton should be allowed an unopposed return. The "600" however, could not accept this policy, in spite of a

[1]Bruce Glasier Correspondence. Glasier to Keir Hardie, 15th July, 1902.

[2]**Ibid.,** Glasier to Keir Hardie, 17th July, 1902.

[3]The Clitheroe Liberals could hardly have chosen anyone less likely to stand. Thomasson's father was a friend of Cadbury, who like the latter had helped to finance "anti-war" candidates in 1900. (See Herbert Gladstone Papers 46047). He had also considered assisting Cadbury in the financing of the *Daily News*. (See A. G. Gardiner, *op. cit.*, p. 211).

speech by Catlow in which he emphasised that they could expect little support from Head Quarters if they decided to fight. An amendment was moved to the Executive recommendation and carried with only about six dissentients. The meeting then unanimously asked the Chairman to be the Liberal Candidate. He refused this invitation but on being pressed, promised to give a final reply within thirty-six hours, although he held out no hope of his answer being any different. Following this the amendment was withdrawn and the original motion was put, only for it to be defeated by an overwhelming majority. In this extraordinary position the meeting came to an end.

Catlow's offer was, of course, a further example of delaying tactics, designed to conciliate temporarily the rank and file of the party, who were spoiling for a fight. However, the indignation was vented against both the national and local Liberal leaders by the local Liberal press. On the other hand the regional and national Liberal press welcomed the now almost inevitable absention as a step in the road to a national agreement between the forces of Liberalism and Labour.[1]

That the new national Liberal policy of conciliating the L.R.C. was at least accepted by the Clitheroe constituency leaders, was demonstrated by the action of Alderman Catlow who on 23rd July refused nomination because : "It would be detrimental to the best interests of the party that Liberalism and Labour should oppose each other in the present election." His decision necessarily involved the abandonment of all hope of a Liberal candidature, "in spite of the extraordinary strength of the Liberal feeling in the constituency" as Sir Ughtred said in a letter to Herbert Gladstone. Although he sympathised with the "resentment of the high-handed action of the Labour people," the late member for Clitheroe was much relieved to think that there would not be a fight. The Liberal Chief Whip must have shared this relief when on 25th July he spoke at Leeds in support of the Liberal Candidate in a by-election there, and praised the action of the Clitheroe Liberals, concluding with a broad hint to the working men of Leeds "that the

[1]See *Daily News*, 21st July, 1902 and *Manchester Guardian*, 24th July, 1902.

Liberal Party meant to do all they could to give them fair representation."

While the General Council of Liberals was hearing the worst, Shackleton was holding his inaugural meeting in the Salem Schools at Nelson. At this gathering and at three others held in other parts of the constituency during the next nine days, he outlined his policy. He called for the amendment of the Workers' Compensation Act, for a pension for "the soldiers of industry" [1] and for improvement in working class housing. Shackleton expressed his support for female suffrage, the payment of M.P's, the redistribution of constituencies, and the principle of "one man, one vote." He supported the disestablishment of the Church of England and promised to strive for the amendment of the Education Bill though he believed there was no chance of defeating it. He believed that increases in land values due to industrialisation should "belong to the state." When asked: "Would you support any measure having for its object the nationalisation of any industry in possession of trusts owned by private individuals?" the Labour candidate replied that monopolies ought to be nationalised and if trusts became monopolies he would include them in his programme. "He was not as well informed regarding the progress of trusts in England as he was concerning those in America, but if the English Trust became like the American he would do his best to check it." [2] It is hard to find anything novel in Shackleton's policy; he was a Liberal trade-unionist and local councillor who showed signs of having been vaguely influenced by most contemporary socialist ideas. One observer said: "I find him a very mild-spoken gentleman, with ideas very little in advance of the average Liberal, though with a strong leaning towards the temperance party. There was absolutely nothing in his short address to which any reasonable Tory working man could object."

[1] See *Nelson Leader*, 25th July, 1902. "They had found over £200 million during the past two years and surely if they could find that amount, they could find £20 million for such a deserving cause as the old age pension scheme."

[2] The confusion of both questioner and questioned was typical of the general attitude to the trusts.

Shackleton was formally returned on 1st August, 1902. A comparison between him and his predecessor is very revealing. Both men were well suited by virtue of their backgrounds and beliefs to represent Clitheroe at Westminster.[1] Both belonged to the Gladstonian section of the Liberal Party, in which both had been prominent locally. In fact, Kay-Shuttleworth had much more in common with Shackleton than had Stanhope, who as an out-and-out radical and one of the most active "pro-Boers," would have been much less acceptable to Conservative working-men than the Darwen municipal councillor. In 1902, they must have been acquainted for some time and though Sir Ughtred was aggrieved at Labour's intrusion, he was instrumental in smoothing over the obstacles to Liberal acceptance of Shackleton's unopposed return.[2] Where the two men significantly differed was in their social origin, a factor of the profoundest importance. The Harrovian son of a baronet was succeeded by a Rossendale clockmaker's son who had gone to the mill at the age of nine.

Shackleton duly took his seat in the House of Commons on 6th August. Even his formal introduction was the subject of some friction, the original intention being that his sponsors would be Richard Bell and John Burns, but strong resistance to this step from Keir Hardie led to the replacement of Burns

[1]They had taken a very similar stand on the Boer War. Shackleton had been Hon. Sec. of the Darwen War Relief Fund. (See *Colne and Nelson Times*, 10th Oct., 1902). But he had signed a manifesto together with other trade-unionists condemning the war as one "waged by capitalists." (See H. J. Ogden, *The War Against the Dutch Republics*, Manchester, 1901, p. 81). Similarly they were both staunch temperance advocates. On 27th Sept., 1902, the Accrington branch of the Independent Order of Rechabites held a meeting to congratulate Brother Shackleton on his election. The new M.P. said : " . . . only those in public life knew the temptations that were placed in front of public men. He would assure them that on more than one occasion he had to upset the glass in front of him." (See *Colne and Nelson Times*, 3rd Oct., 1902). For Sir Ughtred's attitude see *Nelson Leader*, 4th July, 1902.

[2]In a letter to the author the Hon. Rachel Kay-Shuttleworth, the daughter of the first Lord Shuttleworth, writes : "Mr. Shackleton . . . became an intimate friend and was often at our London house and at Gawthorpe Hall. Many years after I was at a function in Burnley and found myself sitting next to Arthur Henderson . . . He turned to me and said, 'Now I have the opportunity of sending your father a personal message through you . . . it is this. He did a GRAND piece of work when he got OUR DAVID in.' "

by Keir Hardie himself. Thus the two members of the Parliamentary Labour Party introduced the third recruit, a ceremony of the greatest symbolic significance to the I.L.P. leader, though it obviously meant little to Shackleton, who apparently would have been quite willing to have been sponsored by two prominent trade-unionists.

Meanwhile Shackleton's union was at last departing from its neutral attitude towards the L.R.C. At a meeting of the U.T.F.W.A. on 26th July, 1902, a resolution was adopted wishing success to his candidature. A resolution in favour of direct representation, moved by the weavers, was accepted unanimously. The Council was instructed to take a ballot of all members on this question. In December, 1902, the result of this vote showed that 84,154 members were in favour of Parliamentary Representation and 19,856 against, while 79,937 were in favour of a levy for the payment of M.P.s and 19,680 against.[1] As a direct consequence the textile workers at their half-yearly meeting on 31st January, 1903 decided to seek representation in Parliament and to affiliate to the L.R.C.[2] Their delegates attended the annual conference of the L.R.C. at Newcastle in February, 1903, when Shackleton was elected on to the committee.

The accretion of 103,000 textile workers represented by far the greatest single gain which the L.R.C. had made to date. Its total trade-union membership increased at one bound from 744,000 to 847,000; the equivalent additional affiliation fee being £51 10s. 0d. Yet the real victory could not be measured in terms of membership or funds, but in the great prestige which accrued to the L.R.C. by the conversion to the idea of independent labour representation of one of the most powerful and best organised of the "old" unions—the union which the Webbs had especially extolled as the superb example of "business unionism."

Thus labour was becoming a force to be reckoned with, a development which was not lost upon either of the political parties.

[1]See Annual Report of the Amalgamated Association of Operative Cotton Spinners, Dec., 1902. 76% of the spinners and 83% of the weavers were in favour.

[2]*Cotton Factory Times*, 6th Feb., 1903.

However, the exigencies of national politics were not always appreciated by local parties. Though it had been the first constituency which had been successfully exhorted to make the sacrifice in the interests of Labour-Liberal unity, Clitheroe was no exception to this rule. Stanhope wrote to Herbert Gladstone of "the heart-burnings amongst Liberals" in the division.[1] At the half-yearly meeting of the Federation of Lancashire Liberal Clubs on 8th October, Alderman Catlow in admitting that the local Executive had acquiesced in the decision of the head of the Liberal Party, complained that the Labour party regarded the Liberal olive branch as a sign of weakness.[2] No doubt this Liberal alarm was accentuated by the buoyancy of the local Labour party, which at the Nelson local elections, in conjunction with the Socialists polled 106 more votes than the Liberal and Conservative candidates combined. Finally at Colne, in the town's only electoral battle labour won a seat.[3] This consolidation of labour strength was to enfeeble the Liberal organization and to make the two towns Labour strongholds for the next half century.

The Clitheroe by-election, then, invigorated the political Labour movement, primarily in North-East Lancashire, but to a lesser extent in all the constituency L.R.C's. It provided a model for Woolwich and Barnard Castle in 1903, when Will Crooks and Arthur Henderson were returned, disproving as it did that Labour could only obtain seats in double member constituencies in conjunction with Liberals.[4] It symbolised the reply of trade-unionism, even moderate unionism, to the Taff Vale decision, a reply which undoubtedly consolidated the unions in their hitherto half-hearted and casual attachment to the L.R.C. Henceforward, especially after the foundation of a fund by which the unions would help to finance parliamentary candidates at the L.R.C. Conference in February, 1903, the unions were to play an ever-increasing part in the activities of the party.

[1]Herbert Gladstone Papers 46051.

[2]*Nelson Leader*, 24th Oct., 1902.

[3]*Colne and Nelson Times*, 7th Nov., 1902.

[4]However, it must be remembered that by-elections provide opportunities for special interests to exert pressure. See D. E. Butler, *The Electoral System in Britain*, Oxford, 1953, p. 166.

Thus the grand strategy of the I.L.P. which had led to the formation of the L.R.C., that is a policy deliberately shaped to convert the unions to independent labour representation, was triumphantly vindicated. One notable result of Clitheroe was that an "old" unionist and early opponent of labour representation, David Holmes, raised no public voice against this new departure.[1] However, the success of the I.L.P's tactics was to be the cause of much heart-searching among I.L.P.-ers,[2] who in the years ahead were to feel that their socialism was being compromised by the expediency of their association with trade-unionists whose policies were more often fashioned to suit the contingencies of the moment rather than the need to provide for the socialist society of the future. The official leadership of the I.L.P. was to undergo serious criticism for such deals as that at Clitheroe. There a Liberal cotton-weaver, whose economic and social philosophy was only a slight modification of John Bright's, was brought as a candidate into almost the only area in Lancashire where the I.L.P. had made any great headway. Above everything else this by-election revealed the extent to which Keir Hardie and MacDonald were prepared to go in order to reconcile the trade unions to labour representation.

By 1906 the influx of trade-union M.P's had produced a Parliamentary Labour Party in which the industrial representatives greatly outweighed the socialists. In fact the addition of Shackleton to the Party meant that Keir Hardie was in a minority, while the adhesion of Bell to the L.R.C. was always in doubt, his conduct eventually leading to expulsion from the party in 1905 due to his co-operation with the Liberals. The interpretation which Shackleton placed on his own "independence" was one which could not have been acceptable to Hardie, based as it was on a conception of the Labour Party

[1]See *Burnley Express*, 23rd July, 1902. "Mr. David Holmes, J.P.was we believe, at first inclined to advocate Mr. Stanhope's claims, but when Labour . . . refused to make terms with Liberalism, even Mr. Holmes veered round."

[2]See *Labour Leader*, 7th Feb., 1903, for a letter from S. G. Hobson, a militant Socialist. "To earnest members of the I.L.P. the term Independent Labour Representation means Socialist representation . . . For myself I have utterly failed to find any special cause for rejoicing over the Clitheroe election . . . "

as a pressure group, which had merely moved from the lobbies to the benches, involving only an extension of conventional trade-union methods and not the creation of a new party.[1] In fact, Shackleton's declared independence was due to an imperative of Lancashire politics, that of not offending Conservative trade-unionists, who though small in numbers in the Clitheroe Division were numerous among the textile operatives in defence of whose interests Shackleton was putting himself forward.[2] It is small wonder then that Shackleton's only difference from the "Lib-Lab's," was that in common with Keir Hardie, he belonged to another generation than men in their sixties and seventies, who had for a decade or more been used to obeying the Liberal Whip. Labour questions apart, Shackleton almost invariably voted in the Liberal lobby, and, as his almost instantaneous association with Dilke and the Radicals shows, he was quite willing to attend joint meetings of the T.U.C. Parliamentary Committee, with Liberal-Labour men and even Radicals, in order to discuss trade-union legislation.

In fact Shackleton's attitude to parliamentary representation was rather confused. Before his election he had made it clear that he regarded his "independence" as a consequence of his union's members agreeing to provide his salary.[3] Trade-union matters, as distinguished from "political questions," were outside the province of both major parties. When it came to "political questions" the Labour Party had allowed him to use his own judgment. He would be : "A representative of the constituency and not of the particular section that returned him." [4] Yet shortly after he was returned, he remarked :

[1]See *Labour Leader*, 26th July, 1902. "While I was selected solely as a representative of Labour it was fully understood that upon matters outside purely trade questions I should be free to use my own discretion. My experience of trade organisations teaches me that it is fatal to associate them in any way with party politics and that it is impossible to obtain direct Labour Representation by working along party lines."

[2]See *Daily News*, 21st July, 1902. " . . . Labour in Lancashire is largely Tory. To have associated their candidate by name with Liberalism would have endangered his hold upon the sympathies of the Tory section of the operative class and that would have weakened his position as a representative of Labour."

[3]*Nelson Leader*, 25th July, 1902.

[4]*Ibid.*

"He was not aware of any question that was not really a Labour question." [1] In fact people interpreted Shackleton's situation in the way they wished. The cotton employers' journal insisted on considering him as a representative of the cotton industry as well as of labour,[2] while quite naturally those sections of public opinion which thought in terms of ideas disapproved of the principle of direct class representation. The S.D.F. weekly gloomily prophesied that: "Men of Mr. Shackleton's type are just the men the wire-pullers of capitalist parties will welcome"; [3] while *The Speaker* protested against "the suggestion of the *Daily News* that Mr. Stanhope should withdraw merely because a working-man has appeared." [4] Later it complained of Shackleton, "on the momentous issues of Imperialism and National Expenditure he is silent" and deployed "the new doctrine . . . that organised interests are to supersede organised opinions as the party forces of the country." [5]

Hardly a "new doctrine," this concept was very prevalent in 1902 and coupled with the seemingly inevitable break-up of the Liberal party in the first half of the year, helped to confuse discussions on the future of the British party system. Keir Hardie had welcomed the development of a "group" system in a House of Commons,[6] in which a small compact Labour group would operate successfully. It did not at that date occur to anyone that "Labour" was a far larger "group" in the country than any other. Even after the election of Shackleton there was no acknowledgment of this fact save by the journal aptly named *New Age*, which on 7th August, 1902, clairvoyantly declared of the Clitheroe result: "We can recall nothing more

[1] It is doubtful if this remark signalised the conversion of Shackleton to socialism. More probably it showed a growing acquaintance with the ideological terminology of the I.L.P.

[2] *Textile Mercury*, 2nd Aug., 1902. "When he gets into the House he will not find himself alone; other weavers are there already; the two members for Blackburn and one for Oldham may be named. Let us hope that where justice and the interests of the trade are concerned they will be found voting together."

[3] *Justice*, 9th Aug., 1902.

[4] *The Speaker*, 12th July, 1902.

[5] *Ibid.*, 26th July, 1902.

[6] See *Labour Leader*, 15th Feb., 1902. "The party system of government is going to pieces."

significant than this in modern political history. We have not heard the last of this we may be sure. Suppose the same tactics are followed in one hundred or two hundred other constituencies? There were at least two or three hundred constituencies in which the same thing could be done as easily as it was at Clitheroe . . . " The end of the Boer War and Joseph Chamberlain's advocacy of Protection healed the breaches in the Liberal Party, so that Keir Hardie's dream did not materialise and the Labour party was forced to become a satellite of one of the major parties until 1918, but the issues which had been debated in 1902 remained alive in all those constituencies where there was the same triangle of political forces as at Clitheroe.

Frank Bealey

University College of North Staffordshire,
Keele.

Games, Capitalism and General Location Theory*

I.

Introduction.

This paper describes conceptual frameworks of location theory. It explains from the perspective of capitalism the capacity and the limitations of many kinds of theories of plant location. It determines basic postulates of location theory and ascertains the changes in these theoretical origin points from one school to another. The examination requires outline of the location models of von Thünen, Weber and other writers. It involves determination of the lines along which there exists a general capitalistic theory of location.

Very few of the writers discussed in this paper sought to derive a general theory of location. Except for Weber and Lösch, the main interest of each appears to be with one kind of location or location influence. Furthermore, only Lösch was concerned primarily with a general capitalistic theory of location. However, even in this case, the focus was formed more by the desire to present a rational system than to describe actuality. As a consequence of its perspective, the assignment of writers in this paper to different niches or schools along the path of a general capitalistic theory of location places them automatically in the position where their respective interests leave gaps compared to the present norm. This position of automatic contrast becomes especially evident in those cases where the game-like uncertainty that arises in capitalism is found to be excluded by the author's assumptions, either implicitly or by his interpretation. It deserves therefore final emphasis to note again the fact that this paper examines particular location models in light of the uncertain game competition that prevails in the private enterprise system. Its

*The author wishes to express his appreciation to the Board of Editors of the University of North Carolina Press for permitting him to draw freely in this article from certain materials contained in his book, *Plant Location In Theory And In Practise* (Chapel Hill : 1956). Also, he wants to acknowledge his thanks to Professors Edward H. Chamberlin and William Miller for their criticisms of an earlier draft of this paper and their suggestions for its improvement.

designation of the elements which would describe the general pattern of location in a capitalistic society (in comparison for example to the location forces of a communistic state) provides at the same moment a base for contrasts between the several conceptual frameworks of location theory that now prevail.

II.

The Least-Cost Theory of Plant Location.

Interest in the theory of plant location may be attributed to three Germans : Launhardt,[1] von Thünen,[2] and Weber.[3] Launhardt is the least known in the United States ; in fact, his influence on American writers is strictly indirect, for his book is not generally available in that country. Launhardt, from many accounts,[4] explained the location of industry as the decision resulting from two variables ; namely, differences in cost and demand at alternative locations.

The second writer, von Thünen, was concerned primarily with agricultural locations. His theory, though designed to explain the type of crops that would be grown at places of varying distances from the market, is applicable nevertheless to manufacturing locations. The location decision rests upon the differences in the cost ot a given crop (commodity) at alternative sites. In turn, the cost differences are due to the land rent and the transportation expense.

Von Thünen's theory can be shown by a simple representation. Assume in Figure 1 that OA is the production cost of a dollars worth of potatoes and $A'S$ ($A''T$) is the cost of transporting the potatoes over a distance of OJ (OK) miles. OB represents the production cost of a dollars worth of wheat while $B'M$ ($B''N$) represents the freight charge over a distance of OX' (OX) miles. Clearly the freight rate is higher on

[1] *Mathematische Begründung der Volkswirtschaftslechre* (Leipzig : B. G. Teubner, 1885).

[2] *Der Isolierte Staat In Beziehung auf Landwirtschaft und Nationalökonomie*, 3rd ed. (Berlin : Schumacher–Zarchlin, 1875).

[3] C. J. Freidrich (trans), *Alfred Weber's Theory of Location* (Chicago : University of Chicago Press, 1928).

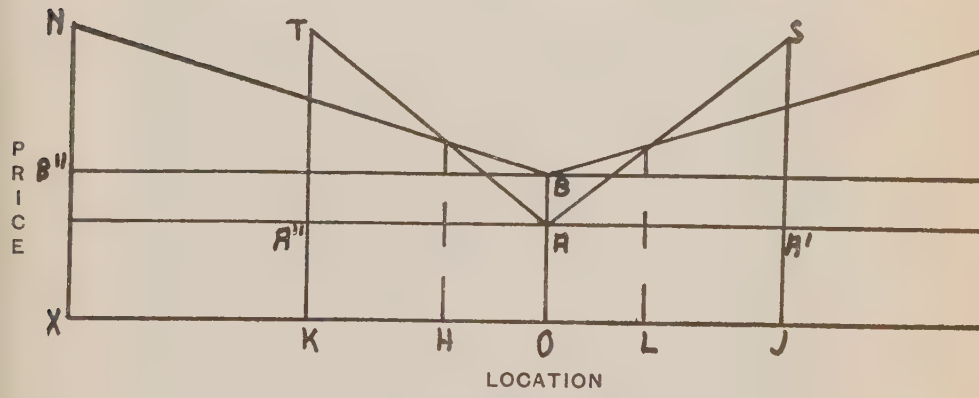
[4] Tord Palander, *Beiträge zur Standortstheorie* (Uppsala : Almqvist und Vicksell Boktrycheri, z.-b. 1935), pp. 139, 140, 231, 247. Erick Schneider, "Bemerkungen zu einer Theorie der Raumwortschaft," *Econometrica*, III, 1935, pp. 79–101, 86ff. August Lösch, *Die räumliche Ordnung der Wirtschaft* (Jena : Gustav Fischer, 1944), p. 6.

potatoes than on wheat. (Cf. the slopes of AS and AT with BM and BN).

Von Thünen's assumption of a uniform homogeneous plain signifies that labor and capital are equal in unit rate and productivity at all locations and that the cost of production (exclusive of transport cost) is everywhere the same. The land rent and the cost of transporting the goods thereby become the effective co-determinants of location. The producers of potatoes will be found in the OL (OH) region, while wheat will be grown between LX' (HX).[1]

Weber's theory of location is procedurally the opposite of von Thünen's. In the earlier writer's scheme, the location is given and the type of production is to be determined; in Weber's theory, the branch of industry is given and the place of location sought. Von Thünen assumes a homogeneous land surface and one consuming center; Weber assumes uneven deposits of fuel and other raw materials and several consuming centers. In practise, however, Weber confines his geometrical representations and general discussions to a given buying point.

FIGURE 1



[1]Cf Lösch, *op. cit.*, pp. 25–39. Palander, *op. cit.*, pp. 71–103. Edgar M. Hoover, *Location Theory and the Shoe and Leather Industry* (Cambridge: Harvard University Press, 1937), pp. 30–33. Also see Arthur H. Leigh, "Von Thünen's Theory of Distribution and the Advent of Marginal Analysis," *Journal of Political Economy*, LIV, 1946, pp. 481–502, 481. And see von Thünen, *op. cit.*, Part 1. Note: von Thünen relaxes his assumption of equal wage payments in the third part of his treatise. This change in assumption permits consideration of the effects of varying labor costs on location.

Weber's approach is also from the side of costs. In his theory, the entrepreneur selects the plant site which minimizes the sum of all expenditures. The locational objective thus involves an optimum in substitution between the factors of cost.

There are three general determinants of location in Weber's theory : transportation, labor, and the agglomerating (deglomerating) forces. Transportation and labor are regarded as general regional factors ; the agglomerating force is classified as a general local factor. The cost of transfer tends to draw industry to the site of least transfer expense, while the cost of labor may cause displacement from that site to a place where the savings in labor cost are larger than the additional transportation cost. The decentralizing tendencies of these two factors are counteracted or intensified by Weber's third consideration : the agglomerating force (marketing advantages, proximity to auxiliary industries, economies of scale), and its corollary, the deglomerating influence (land rent). These considerations draw industry closer together or disperse it, depending upon the respective strength of each force.

Weber's theory, like von Thünen's, involves substitution between transport cost and non-transport cost factors. Conceptually, a series of isodapanes (points of equal transfer costs) are conceived as placed around the minimum transfer cost point. The curve farthest away from the minimum isodapane represents sites of highest transfer burdens, *etc.* The isodapane which exceeds the least cost transfer point by an amount equal to the maximum non-transfer cost economies that are obtainable at an alternative site is the critical isodapane. If the optimum production (non-transfer) cost site lies inside the critical curve, the production (non-transfer cost) advantage is worth more than the difference in transfer cost. If it lies outside the critical isodapane, the economy in production is less than the extra transfer cost that would be entailed in the shift.[1] Obviously each location has certain transfer cost advantages or disadvantages when compared to other locations and similarly non-transfer cost advantages or disadvantages.

[1]For a discussion of contour systems, see Palander, *op. cit.*, p. 212 ; Lösch, *op. cit.*, p. 12 ; and Hoover, *Location Theory*, *op. cit.*, pp. 32–55.

Weber's analytical framework is the purely competitive one. All buyers are assumed to be concentrated at given buying centers and each seller locates in respect to and has access to some buying center. Each seller has an unlimited market, for price is a datum and demand for the output of a firm is infinite relative to its supply. There are therefore no monopoly gains derived from location. Weber not only follows this *type of* abstraction from demand (that is, an assumption of a market type in which the demand factor does not exercise an active role in determining locations), but, further, he disregards institutional costs such as taxes and interest in his theory. These omissions arise from his desire to formulate a general theory of location which is applicable to all economic systems.[1]

Weber's influence was universal. There arose a school of writers who sought to explain manufacturing locations as dependent upon the struggle to locate at places at least cost. While some of these writers (Palander [2] and Hoover [3]) expressed interest in the size of the market area, and, therefore, in a sense, were concerned with the demand factor, the majority disregarded demand in the von Thünen-Weber tradition. Thus Predöhl was interested solely in developing a substitution cost analysis ; [4] Ritschl inquired historically into

[1]Interest, insurance, taxes, climate, and management are all excluded from Weber's system, either because they are institutional in origin and/or of such special nature that they influence only a few plant locations.

[2]*Op. cit., passim.* And see p. 221 where Palander in discussing von Thünen and Weber stated that the mutual connection between market and location has been neglected.

[3]*Location Theory, op. cit.*, see in particular Chapter III. Also see Hoover, *Location of Economic Activity*, 1st ed. (New York : McGraw Hill, 1948), Chapter IV.

[4]A. Predöhl, "Das Standortsproblem in der Wirtschaftstheorie," *Welt Wirtschaftliches Archiv*, April, 1925, pp. 294–321. Also see, G. Cassel, *Theoretische Nationalöhonomie*, 3rd ed. (Leipzig : W. Schall, 1923), pp. 83–96. Witold Krzyzanowski, "Review of the Literature of the Location of Industries," *Journal of Political Economy*, XXXV, 1927, pp. 278–291, 288, for a discussion of Cassel. O. Engländer, "Kritisches und Positives zu einer allgemeinen reinen Lehre vom Standort," *Zeitschrift für Volkswirtschaft und Sozialpolitik*, V, 1927, pp. 435–505. A. Predöhl, "Zur Frage einer allgemeinen Standortstheorie," *Zeitschrift für Volkswirtschaft und Sozialpolitik*, V, 1927, pp. 756–763. W. Isard, "Distance Inputs and the Space-Economy ; Part I : The Conceptual Framework," *Quarterly Journal of Economics*, LXV, 1951, pp. 181–198, 195–197.

the changing pattern of cost and location ; [1] Link [2] and other students of Weber attempted to measure the displacement of certain plants from the least cost transportation point, which displacements were explained as the result of labor and agglomerative influences ; Holmes and other writers discussed industries in terms of orientation to materials, labor market, and other cost factors ; [3] Dechesnes [4] was similarly inclined.[5]

The significant features of the von Thünen-Weber type of theory sum up to : (1) the emphasis on the search for the least cost site ; (2) the assumption of a given demand at a particular buying point which remains unaffected by the locational selection of the seller ; and, (3) the inherent disregard of the game-like locational interdependence of firms. Each of these three main aspects points out the basic limitations of this school of thought, for, (1) site-selection (including purely personal site-selections) involves substitution among not only the cost factors at alternative locations, but also the demand factors at the different sites, (2) market demand (and the utility gained by the individual) is a variable, significantly affected by, and affecting, (3) the locational interdependence of firms (and of people). It is with the second (2) and third (3) objections to the von Thünen-Weber type of theory that the following section is concerned ; an analysis of the initial complaint is reserved for section IV.

[1] "Reine und historische Dynamik des Standortes der Erzeugungszweige," *Schmollers Jahrbuch für Gesetzgebung*, LI, 1927, pp. 813–870.

[2] *Die Lederindustrie* (*Erzeugende und Verarbeitende*) (Tübingen, n.p., 1913). See Hoover, *Location Theory*, *op. cit.*, p. 119.

[3] See E. Holmes, *Plant Location*, 1st ed. (New York : McGraw, Hill 1930). See also the empirical type of studies of E. Hedlund, *The Transportation Economics of the Soybean Processing Industry* (Urband : University of Illinois Press, 1952) and D. C. Hague and P. K. Newman, *Costs In Alternative Locations* : *The Clothing Industry*, National Institute of Economic and Social Research (London : Cambridge University Press, 1952).

[4] *La Localization Des Diverses Productions* (Bruxelles : Les Editions Comptables, Commerciales et Financieres, 1945).

[5] The cost factors described by recent writers are not identical to those that were considered by Weber. Weber's original interest lay in the evolution of a theory that would be applicable to all economic systems. Other writers (Holmes–Hoover–Palander–Dechesnes) stress the private capitalistic economy and thereby include institution factors within their framework. The ultimate objective of these treatises is the formulation of a system which includes all factors of location in a private capitalistic economy.

III.

The Interdependence Theory of Plant Location.

The von Thünen-Weber theory is quite applicable to the problem of agricultural locations, and to certain manufacturing locations, especially those wherein small scale manufacturers sell to large size wholesalers. But, the increasing awareness in recent years of the limitations of the von Thünen-Weber purely competitive framework led to another approach to the problem of plant location. Under the influence of Fetter,[1] Hotelling,[2] Lerner and Singer,[3] Smithies,[4] Chamberlin,[5] and other writers, interest centered upon locational interdependence. This analysis was predicated upon the theory of duopoly; it abstracted from cost [6] and explained the location of firms as the endeavour to control the largest market area. The methodology and conclusions were in contrast with Weberian findings, but the problem was essentially the same.

In the Weberian framework, price is given and buyers are assumed to be concentrated at point-formed consuming centers. Because there are not any sales-price advantages that can be gained at alternative locations tributary to the selected market, the location decision in reference to a given market center rests solely upon cost. In the Fetter-Hotelling framework, delivered price varies with location, for buyers are assumed to be scattered over an area. Each firm (subject to identical costs and quoting the same f.o.b. net-mill price) is able to sell to proximate buyers at delivered prices which are lower than those of their rivals. This industrial pattern signifies the exclusion of competitors from the market area that surrounds a seller's plant. It makes the location of each seller dependent upon the number of consumers he can monopolize at the different

[1] "The Economic Law of Market Areas," *Quarterly Journal of Economics*, XXXVIII, 1924, pp. 520–529.

[2] "Stability in Competition," *Economic Journal*, XXXIX, 1929, pp. 41–57.

[3] "Some Notes on Duopoly and Spatial Competition," *Journal of Political Economy*, XLV, 1937, pp. 445–486.

[4] "Optimum Location in Spatial Competition," *Journal of Political Economy*, XLIX, 1941, pp. 423–439.

[5] *The Theory of Monopolistic Competition*, 5th ed. (Cambridge: Harvard University Press, 1946). See Appendix C.

[6] These theories assumed that the cost of procuring the raw materials and processing them were equal at all locations.

locations. The attempt to control the largest number of buyers (market area) at prices which yield greatest returns is the driving force behind orientations to market areas. Under this framework, the size of the market area that belongs to a firm is determined by the locational interdependence of the firms.[1]

The assumption of a market area leads the writers on locational interdependence into an appraisal of the influence of (A) the shape of the industrial demand curve, (B) the shape of the marginal cost curves, and (C) the height of the freight rate on the location of industry. These factors influence the conjectural hypotheses of entrepreneurs about the location policies of rivals ; they, along with others, such as the historical locational competitiveness of an industry, determine the degree of dispersion when conditions of sale are perfectly competitive except as regards space and number of competitors. The impacts on theory of these types of forces are shown in the following models.

Assume :

1. Buyers are uniformly distributed over a linear market. (Conception of buyers as distributed along a line instead of over an area simplifies the problem without affecting the nature of the conclusions).

2. Buyers are indifferent to sellers ; neither sellers nor product are differentiatied, except for location.

3. Sellers are indifferent to buyers ; all buyers are homogeneous in every respect, except location.

4. The seller's procurement and production costs are everywhere equal. Marginal costs for competitors are constant. With no extra great loss in realism, average and marginal costs are assumed for the present to be not only equal but zero.

5. Each competitor sells on a non-discriminatory f.o.b. mill basis ; that is, each buyer pays the same net-mill price.

6. Each competitor is capable of supplying the entire market.

[1]Each seller becomes a locational monopolist when consumers are scattered and firms practise f.o.b. selling. In Weber's system, sellers compete actively for the same market, as in perfect competition.

7. Cost of transport is at the same rate per unit of distance throughout the market area.

8. Each competitor is free to move his location instantaneously and without cost.

9. The security motive does not influence the locational choice.[1]

10. Each competitor charges the same net-mill price.

(A) If the demand for the product is infinitely inelastic (each buyer purchases one unit of product at one unit of time), duopolists will concentrate together at the center of the market area. From this vantage point each can provide the remotest extremities of the market while limiting the rival to no more than his proportionate share of the market.[2] On the other hand, an infinitely elastic demand disperses the sellers to the point where they locate in perfect harmony with the location of buyers.[3]

If the demand curve for the product of an industry is assumed to be of the negative sloping straight line variety,

[1]It is readily understandable that the security motive may determine the location and not maximum profits. For a discussion of the security motive, see K. W. Rothschild, "Price Theory and Oligopoly," *Economic Journal*, LVIII, 1947, pp. 299–320. And see G. Tintner, "The Theory of Choice under Subjective Risk and Uncertainty," *Econometrica*, IX (1941), pp. 298–304. Of course, F. Machlup suggests in *The Economics Of Seller's Competition* (Baltimore: Johns Hopkins University Press, 1952), pp. 51–56, that the security motive is part and parcel of maximum profits. But distinction between the two does have practical as well as theoretic basis. See cases 5 and 6, M. Greenhut, "Observation of Motives to Industry Location," *Southern Economic Journal*, XVIII, 1951, pp. 225–228.

[2]This is the celebrated Hotelling case. See H. Hotelling, *op. cit.*, p. 52. From any firm's standpoint, the sharper competition with rivals, due to proximity, is offset by the greater number of buyers over whom it has an advantage. As we will see later on (note 2, page 72), the *general* conclusion for this case (when there are more than two firms) is dispersion.

[3]See Morris A. Copeland, "Competing Products and Monopolistic Competition," *Quarterly Journal of Economics*, LV, 1940–1941, pp. 1–35, 3, 23, where an infinitely elastic demand curve is assumed, and in the absence of freight absorption, it is concluded that sellers will disperse in place and numerical accordance with the dispersion of buyers. And see A. P. Lerner and H. W. Singer, *op. cit.*, pp. 145–186. The writers assume an inelastic demand over a range extending from zero to a finite upper limit. And " . . . on the assumption that *A* takes *B's* position as given (*i.e.*, location and price), he will not behave as in Hotelling's hypothesis, but will keep at a distance or attempt to drive him out," p. 154.

firms will not localize at the central locations nor will there be any great assurance of their perfect dispersion. Only in the long-run would they tend definitely to locate as the multi-plant monopolist, who, in case of two plants, locates one at the first quartile and the other at the third quartile of the linear market, and in the case of three plants locates each at the first, third, and fifth sextiles, *etc.* This scattering minimizes freight costs and maximizes profits. Its necessary precondition is the entrepreneurial conjecture that any attempted incursion in market area, either by way of lower price or locational movement to the center, will be met by the other competitor. This means that rivals must have knowledge that the quartile-sextile, *etc.* type of location is most profitable for each firm. In turn, this understanding is determined by several factors, among which the shape (straight line, concave, convex) of the demand curve takes a leading role.

When the demand curve is convex to origin, the fact that symmetrical locations are most profitable is not quite so obvious as it is in the case of a negatively sloping straight line demand curve. Demand curve convexity causes an added element of doubt regarding the rival's knowledge of which location maximizes profits. Convexity to origin signifies that absorption of freight costs in net-mill price will be great and that sales to the peripheries of the market area may be gained at locations which are central to the entire area. Movement toward quartile locations (though advantageous) is shrouded in doubt. Conversely, when the demand curve is concave looking downward, net-mill price may be greater on account of freight than the net-mill price that would have prevailed in the absence of freight.[1] This condition promotes the quartile-sextile-*etc.* type of locations as a firm may expect a competitor to seek distant sites in order to gain sales at the periphery of the market area.

[1]This result can be understood simply by realizing that the more concave the demand curve, the further to the left will be the intersection of marginal revenue with a marginal cost curve that has been increased because of freight costs. The delivered price that is shown by this intersection will be higher than the net-mill price (which did not include freight cost in its consideration) by more than the freight burden that exists.

(B) The shape of the marginal cost curves is a similar determinant of locational interdependence (concentration or dispersion) when buyers are assumed to be scattered over an area rather than concentrated at a point. If marginal costs increase with expanded output, freight absorption in net-mill price is great. It follows that the impact of freight cost on price may be so slight compared to the price that would exist without spatial effects that firms at a central location may well be able to supply the peripheries of the market. To the extent that sellers seek large sales radii, doubt therefore exists regarding movement of a rival to distant sites. Conversely, when marginal cost curves are declining, freight absorption in net-mill price is slight ; in fact, net-mill price may actually increase because of freight cost.[1] It follows that there is greater assurance in this case of movement by a rival away from the center of the market.

(C) The height of the freight rate is a locational determinant which influences the cluster or dispersion of industry. Low freight rates (or any lowering of freight rates) leads to localization in the center as the hinterlands are more serviceable from a distance. On the other hand, high transit cost (or increase therein) tends to disperse the particular industry.[2]

There has been dispute in the literature concerning location of three or more firms under inelastic demand. Hotelling [3] offered the view that sellers would localize at the center of the market. But, this view has only slight merit under game

[1]If a linear marginal cost curve is negatively sloping at the same degree as average revenue, then the intersection between a marginal cost curve of this type, which furthermore has freight cost added to it, and the marginal revenue curve that is associated with the defined average revenue curve will leave a delivered price that is greater *by the amount of the freight* than the total price that would have existed in absence of freight costs. In this case there is no freight absorption at all, and the net-mill price with and without freight cost considerations is the same. On the other hand, if the marginal cost curve is sloping negatively at a steeper rate, or at a more moderate rate than average revenue, the net mill price that includes consideration of freight costs will be greater, or less than, the nonspatially formed net-mill price, as the case may be.

[2]And see Schneider, *op. cit.*, p. 80, wherein the further conclusion is drawn that, *cet. par.*, "die Grösse des Absatzgebietes . . . dem Quadrate des Frachtsatzes umgekehrt proportional ist."

[3]*Op. cit.*

theory; and while it could happen that firms would concentrate,[1] probabilities suggest some immediate dispersion.[2] According to Lerner and Singer,[3] the dispersal influences which exist are formed by the need of each firm not to be caught between the other two, that is unless the distance between each firm are either substantial or its opposite, microscopically small. To these writers, the locations are unstable over time. The

[1]If plans must be made well in advance of any construction, so that land options, and contracts to purchase materials, to float securities and to borrow funds, when once arranged can be ill withdrawn, the location problem under uncertain conditions becomes like most games; it involves hidden play. Only at a given moment does revelation of decision come about; and then, the revelation by one is followed closely by that of another. This high simultaneity in announcement of decision indicates that rivals had to conject and plan without precise knowledge of each others ideas. The uncertain conjectures of a firm forms the basis for the theory of games.
On *a priori* grounds it would appear that the conjectures of any one firm must be that its two other rivals, operating under the same economic climate and subject to comparable expectations, will plan each to locate at a substantial distance from the other or else intend to locate close together. The numerical values to be assigned to these possibilities must exceed those accorded to the chances for opposite plans by each firm. Considering then only the cases of comparable plans, it may be that each firm with plans for dispersion will locate at the opposite quartile, or else of course their plans may not have succeeded and both may locate accidently at the first or third quartile together. On the other hand, similarity in plans may suggest intent to locate together; in this event, planned location at the market center is most probable. Manifestly, it is only in this latter case that the third firm will do well in selecting a site distant from the market center. In the other cases, central location suggests advantage. Because there are three ways to one in favor of central location, the minimax site becomes this one. Of course, any firm may then carry the thought that if the other two act according to this minimax site, its own minimax place would become the distant location rather than the central one; but comparably, this same extra thought may be entertained by the others and hence would be discarded. Further, let it be noted that because accidental concentration of all three at the first or third quartile would encourage numerical expansion of the industry, the firms would tend to locate together near the center rather than at the quartile type of site.

[2]Any tendency to localize fully at the market center is weakened by the fear of being hemmed in between the other two firms. Thus any one firm will plan location near the center but somewhat to the right or left of center. Accidentally, the Hotelling solution of full concentration may arise. But chances are that two firms will be to the right (or left) of center with the other firm sited to the left (or right) of center. Indeed, if the ends of the line are joined to form a circle, probable separation of all firms is sharply revealed. (*i.e.*, see Edward H. Chamberlin, "The Product as An Economic Variable," *Quarterly Journal of Economics*, LXVII, 1953, pp. 1–29, 19, 20).

[3]*Op. cit.*, see their discussion, pp. 176–179.

cycle may find initially three firms at the center, then two at one quartile, and one at the other, then all three back in the center, *etc.* The underlying force behind the wave of relocations is the assumption by each competitor that the site-selections of the others are fixed. Chamberlin,[1] on the other hand, deduces the possibility that two firms may be located quartiles and the third competitor someplace between them. He offers the opinion that the continual shifting of sellers seeking advantage carries the dispersion at least this far and perhaps further.

For cases of more than three firms, a clustering form of dispersion as in the Lerner-Singer adoptions may arise, or else a more complete dispersion in the Chamberlinean tradition appears likely.[2] Significantly, the tendency toward scatter becomes stronger with greater number of firms. It remains always possible, however, that the great fear of being hemmed in may stimulate tacit agreement to localize together at the market center in sharply concentrated form. Obviously, such extreme localization would be somewhat unusual as well as quite unstable.

Once the assumption of an infinitely inelastic demand curve is dropped, it becomes apparent that the firms will disperse in a more ideal way. Such locations follow the limitation set by a finite demand. Indeed, if in addition to demand elasticity the cost of transport and the marginal costs of production are assumed to be high, the scatteration of firms will proceed quickly in more perfect conformance with location of buyers. Each firm would be more willing to announce its plans for distant location without waiting to play the game of cat and mouse.

While economists generally have postulated equality of production costs everywhere, the above analysis can be broadened to include the case of divergence of production costs at alternative sites. But this type of model may be considered

[1] *The Theory of Monopolistic Competition, op. cit.*, Appendix C.

[2] Over time, on a circular street, or over a circular market area there exists stronger tendency to full dispersion. If the fourth, fifth, and still later firms locate after leaders are already in the market, the dispersion may be complete or in clustering form depending upon the pattern that prevailed prior to the entrance of the later firms.

with greater advantage in a separate section to follow. It remains, at present, to summarize and appraise the theory of locational interdependence.

The general discussion of this section points to the following main conclusions : (1) the tendency to disperse depends upon the height of the freight cost, the elasticity of the demand function, and the characteristics (slopes) of the marginal costs ; these factors, along with historical practise, determine the degree of competition in location. (2) Each seller seeks to control the largest market area. The actual location is determined by the type of interdependence existing between the seller and his rivals. (3) Each seller becomes a spatial monopolist when sellers and buyers are separated geographically from rivals and when the selling industry uses the non-discriminatory f.o.b. price system. (4) Effective demand varies at alternative sites because of freight costs and the *location* of rivals. (5) The theory explained is designed mainly for manufacturing plant locations where raw materials and labor supply are fairly ubiquitous and where demand is the vital variable factor that is within the control of the firm. Also, the overall theory probably has high relevance for locations of functional middlemen and certain special wholesaling and retailing situations. Indeed, it can be introspected that the individual who derives his satisfactions (religious, social, climatic,—including amount of space and air-entertainment, *etc.*) over an area is in a position similar to that of the firm. Like the business establishment, he jockeys for the right location within a landscape of diverse utilities. (6) Three or more firms locate similarly as two. All other things being equal,[1] the force of dispersion is stronger the greater the number of firms.

The limitation of the theory of locational interdependence is patent : by design it abstracts from cost in similar manner to the abstraction from demand of the least-cost theoriests. It yields therefore only a one-sided theory which reveals a special type of explanation of underlying forces of location in space. The following section attempts to avoid the specialization

[1]Example of inclusion in the *ceteris paribus* assumption is the type of product being sold, shopping good or convenience good.

in assumptions of sections II and III. It treats cost and demand as variable factors that govern the selection of sites. This examination uncovers the intense uncertainty that develops over rival strategies.[1]

IV.

The Theory of the Maximum Profit Plant Location.

The future influence of August Lösch on location theory is hard to predict ; yet this writer reached the core of the problem in his recent book : *Die räumliche Ordnung der Wirtschaft.*[2] From Lösch, more perhaps than from any other writer, the theory may be advanced that the location of a manufacturing establishment depends upon the firm's cost of production at alternative locations and the market area which it is able to control from each site. But while he explicitly recognized variability in costs and demand at alternative locations, and emphasized long-run competitive adjustments in location which minimize freight cost, he did not combine in his analysis an investigation of cost with an appraisal of locational interdependence. In fact, Lösch really gave only lip service to the cost factor of location, for he abstracted largely from this determinant in his basic model by assuming a homogenous land

[1]The line problem of location has one other aspect besides its emphasis on locational interdependence. This side come into play under the model of many many firms in the market. In this case, each new participant is forced to select location in accordance with the remaining availability of sites. In turn, the actual number of firms in the market and their dispersion takes form as the central problem for analysis.
Resolution of the problem of number and dispersion depends essentially over time on the factors of convenience in buying and the production economies of scale. These underlying determinants (demand and cost curves) are shown by Lewis to lead toward the ideal number of firms in the market and associatively toward the ideal size of firms when economies of scale are tied up with a quite elastic demand. This tendency is lessened when scale economies are related to a less elastic demand. Space relations are thus shown to be one aspect of monopolistic competition. The tangency solution involves higher than optimum cost. See W. Arthur Lewis, "Competition In Retail Trade," *Economica*, XII, 1945, pp. 202–234.

[2]*Op. cit.*, see in particular Chapters 1–15, and p. 19. To look at the point of lowest cost is as wrong as looking for the point of highest sales. "Jede solche einseitige Orientierung ist falsch. Richtig ist allein die Suche nach dem Ort grössten Gewinns."

surface. Even more vital, he was not concerned with the principles of interdependence that were described in section III.

Briefly, his theory is as follows :

In the long-run, each firm sells over a hexagonic shaped market area, for this type of shape minimizes total distances from its center to all points within the polygon.[1] Because the trading area of each firm is in the form of a hexagon, a system of hexagons (all firms and industries) exists, which, when rotated around a common center, creates a plain (or surface) marked by six sectors that are relatively open (barren), and six sectors that are dotted by concentrations of individuals and industries (full). The advantage of this pattern lies in the inter-industry and inter-consumer advantages of agglomeration ; while, within given industries, the entrepreneurial scattering (and hexagonic market areas) maximize total effective demand for each product. Thus, in Lösch's theory, the agglomerating cost factor promotes inter-industry concentration, but the demand factor leads to a honeycomb type of dispersion of firms within an industry.

While Lösch's theory is highly informative, (1) his non-inclusion of cost differentials, other than those attributable to advantages of agglomeration and transportation, (2) his failure to carry to its real-world result the full impact of agglomerating advantages on the locations of firms that belong to a given industry, (3) his consequent omission to combine an analysis of cost and demand factors in one model, and thus (4) his disregard of the forces described in section III which cause extra-ordinary concentrations of homogeneous business units, render his conclusion (that attempts to maximize effective demand lead to a hexagonic type of intra-industry dispersion of firms) clearly inadequate as a general explanation of plant location in a privately capitalistic economy. This section of the paper embellishes upon the third complaint. It presents a model in which both cost and demand are treated as variables at all locations. The other critiques are left for the fifth and

[1] In generalizing his picture of the ideal size and number of firms, Lewis also deduces the hexagon as the market area shape that yields the ideal stable equilibrium. See *op. cit.*, pp. 205, 206, for a demonstration of and specification of the size of the hexagon that yields the equilibrium.

final section of the paper where the question whether past and present location theory adequately explains site-selections in a privately capitalistic or a purely communistic society is explored.

Suppose the B location (Figure 2—at the center of the market) is the least cost position, and that other sites situated away from B entail heavier procurement and production charges. Where will firm A locate?

If the costs of production and procurement at locations other than B exceed transportation savings from these sites to any buyer, no location other than B is feasible. Figure 2 reveals this situation (the BP and AP' stem include equal mark-ups above cost). As a general rule, it follows that where P_B (Price of B) + T_B (Transportation from B) is $< P_A + T_A$, at any location other than B, all firms must locate at B.

Where the procurement and production costs are significantly larger at other locations, but an untapped segment of the market remains which can be supplied most economically by a plant situated nearby, a location at A in Figure 3 is feasible. In this case, the firm may charge a higher mark-up above cost so as to extend the AP' stem to AP''. High cost locations made in reference to such market segments would

FIGURE 2

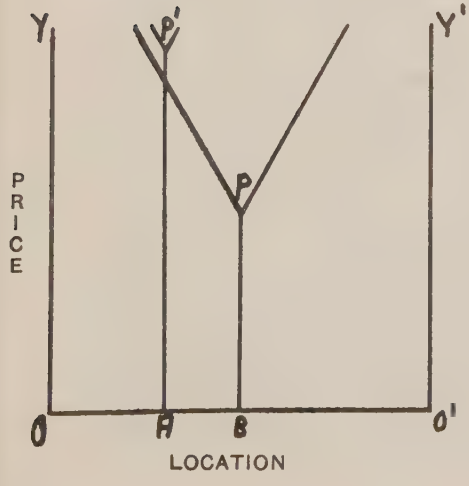
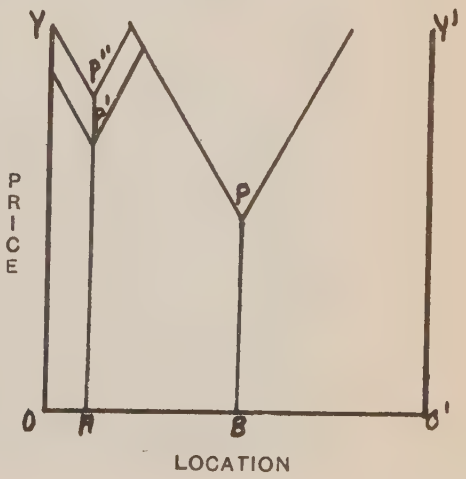


FIGURE 3



generally be adopted by small sized plants.[1] On the other hand, where owners plan larger scale enterprises, they must compete actively at or near B in order to attain this objective (see Figure 4).

Under the assumption of equal costs, it has already been noted that quartile positions result initially, or in the long-run, when each of two firms assume any price set or any location adopted by the one will be met identically by its rival. However, if costs are not identical at all locations, this quartile movement does not necessarily take place. The determining factors are the sales, the net-mill prices, the costs at the two locations,[2] and the conjectural hypotheses of the locators. Significantly, imperfect seller's knowledge of each others costs arise when costs are not uniformly equal over space. This combination introduces a special element that must be considered in theory.

The essential difference between the cases of equal and unequal cost is the element of extra doubt. Though past performances indicate quasi-monopolistic policies, the future may reveal antithetical inclinations. Quasi-monopolistic practices are based upon economic advantages gained by non-actively competing firms. But B may embrace competition in location in this particular market, *if* convinced that he can not lose thereby. Such conviction follows the belief that the high costs at the quartile or other locations can not be made up by savings in freight, the consequently smaller freight absorption, an anticipated lower average delivered price, and the expected increase in sales. If this be the conjecture of the owner of firm B, he has nothing to gain and everything to lose by location at a site other than the central (least-cost) site.

[1]Ritschl., *op. cit.*, pp. 813, 829, 870, 848 . . . Far away sellers can exist even though they have higher production costs and a smaller sales radius. The freight rate protects these firms.

[2]Sales at the quartile location are greater if the delivered price to a large number of buyers is sufficiently lower from this location than the delivered price from the least-cost site. The comparison between delivered prices depends upon the extent to which savings in freight compensate the higher marginal costs. Finally, the net-mill price and associated sales must be so much larger as to raise the net average revenue curve sufficiently in respect to the higher average cost curve. An increase in sales alone does not make it profitable for firms to locate at sites of higher cost.

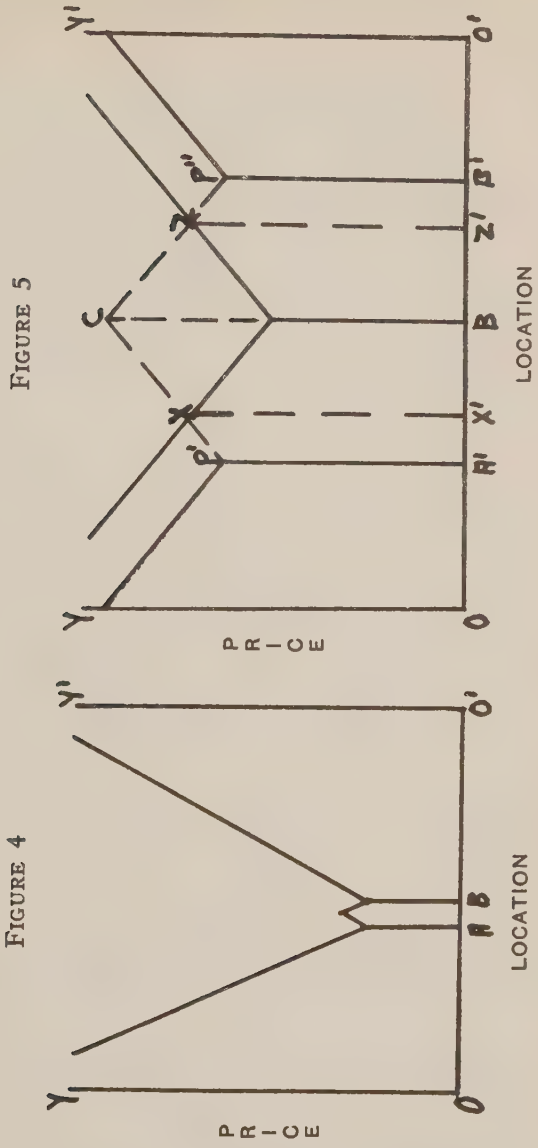


FIGURE 5

FIGURE 4

Conversely, *A* goes directly to the quartile, or other positions, only if convinced that *B* always locates symmetrically.[1] The solution is similar in the long-run.

The case of unequal costs suggests two new principles : (1) quasi-monopolistic tendencies in price are not so closely associated with like practises in location when costs are equal everywhere. The less definite correlation results from the conjectural hypothesis discussed above : the uncertainty of the situation. To this must be added a new consideration, which is also the second principle, (2) the rivalry from substitutable products is probably more pronounced, more influential, and therefore a determinant of location when costs are dissimilar at alternative locations.

The assumption of a least-cost location in respect to the entire market area suggests the possibility that like industries are having similar problems. If it can be presupposed that the least-cost location for the industry in question is also least in cost for many other industries, and particularly for an industry producing similar goods, the problem of substitutable (hetero geneous) products ascends in locational importance. This aspect (location factor) can be seen readily by reference to Figure 5.

In that figure, quartile locations cause an increase in price between *X'* and Z' as compared to the price existing if a central location was adopted. This increase in delivered price may cause extra great losses in the district between *X'* and Z' because buyers shift to the substitutable product. Movement away from the central location (*i.e.* to *A'* and *B'* in Figure 5) leads not only to the normal losses between *X'* and Z' on account of higher delivered prices, but, in addition, extra-losses result from the high cross-elasticity of demand at the

[1]The herioc assumption of zero costs in relocation would be necessary to justify a short-run quartile location.

central site.[1] Furthermore, there is a smaller probability that a firm (industry) producing a substitutable product will locate symmetrically (disperse) with rivals, because such company has formed a custom by differentiating its product. These conditions mean that there is less likelihood that heterogeneous firms will pursue quasi-monopolistic location policies in respect to each other. Manifestly, where firms are large and where costs vary widely among locations, they will concentrate, *ceteris paribus*, in the particular city or district which is least in cost.[2] The location of the large firm is accordingly made with particular emphasis on the least-cost site in relation to the whole market area. These firms do well in emphasizing the location of their rivals. Only the foolhardy dare chance the initial move if doubt exists as to the probability of symmetrical locations.

The location of the smaller firm is somewhat more flexible according to this reasoning. These firms seek smaller market segments, and, therefore, may locate at the least-cost position in reference to a given "particular" segment of the entire market. The smaller firm chooses its site with greater freedom in this respect ; it is more ready to avoid locations near competitors ; *a priori*, smaller plants generally are found in the less industrialized areas.[3]

[1]This general discussion can be understood readily by reference to Figure 5. Respective locations at A' and B' raise the selling price between X' and Z' above that which would exist if the optimum cost location at B was chosen. The more effective utilization of the $X'O$ and $Z'O'$ districts, *ceteris paribus*, may or may not compensate the loss between X' and Z , depending upon which segments of the market had the higher cross-elasticities. To the extent that point B represents a highly industrialized location, the loss to rival industries may be greater than the hinterland gains. On the other hand, where costs are equal everywhere, the problem of substitutable products is most pronounced at the market extremities. This latter result is derived from the fact that delivered prices are highest at places most distant from the seller's factory, and therefore the demand is, generally, most elastic at such sites. Equal costs indicate quartile locations for homogeneous firms ; unequal costs may not suggest the same result.

[2]See P. Sargent Florence, *Investment Location and Size of Plant*, National Institute of Economic and Social Research, Study VII (London : Cambridge University Press, 1938), p. 84. The large firms within an industry concentrate at given points.

[3]This statement, of course, excludes from consideration the service companies. See Florence, *op. cit.*, p. 84, for statistical verification of this deduction. . . . Small plants disperse more than the large firms. They locate generally in the less industrialized areas.

The inclusion of cost and demand factors in one model points out the need for a broader statement of the determinants of plant location than one which concludes that firms seek the location of least cost, or one which holds that firms seek the location offering the largest market area. This need is fulfilled by the concept of the maximum-profit location.[1] By definition, this location relates to that site from which a given number of buyers (whose purchases are required for the greatest possible profits) can be served at the lowest total cost. And while the lowest level of average production cost at this site may be higher than that which exists at alternative ones, the monopolistic control gained over larger numbers of buyers (spread over a market area) makes it the maximum-profit location at the optimum output.[2]

This definition of the locator's objective recognizes the fact that one location may offer lower manufacturing unit costs at a given output than another, but that the relative positions may be reversed as the market area of the firm expands and differentials in transport cost appear. For example, high cost manufacturers who extend their operations by water transportation may gain lower total unit costs to distant buyers than competitors linearly nearer to these buyers who are forced to use the railway or highway networks.

This concept of the maximum-profit location does not have to be confined to an analysis in which the demand or cost factor is held constant. It lends itself readily by definition to an examination of cost and of locational interdependence. When uncertainty prevails, implicit modification of the concept is, however, required. A minimax force arises which finds impropriety in the more daring movement to the quartile. As an implicit part of the concept of the maximum profit location, the minimax principle finds effect in its tendency to draw firms together. In substance, the strategy of games may suggest concentration, when dispersion otherwise would result.

[1]This maximum profit location concept contains many sides. Essentially, it excludes psychic income factors.

[2]Weber's least-cost location fits the above definition in part ; its compliance is limited, however, because Weber's least-cost location must emphasize the cost of procuring, processing, and distributing goods to a *given* buying point and thus it cannot convey spatial implications.

Significantly, the generic maximum profit location theory appears adaptable to not only the general type of manufacturing plant location, but also to locations on other levels, particularly retailing as well as that of the individual.

It is manifest that the described location theory offers insight into an objective of plant location, and the ways and means by which plant managers seek to achieve this objective in their selections of plant sites. It points to the following main economic conclusions : (1) When firms sell to a given buying point, they seek the least-cost location in reference to this consumption center and ignore the locations of rivals in selecting their plant-sites. (2) When firms sell over a market area, their site-selections are influenced greatly by the location of rivals. (3) In selecting a plant site, each firm seeks that place which offers the optimum sales output at a cost that can not be matched elsewhere. (4) When firms sell over a market area, and unequal costs exist at alternative locations, the minimax force of concentration becomes an inherent part of the locational patterns. (5) When firms sell over a market area, the tendency to disperse depends upon the height of the freight cost, the elasticity of the demand function, the characteristics (slopes) of the marginal costs, the degree of competition in location, the degree of competition from substitutable products at the various locations, and the homogeneity or heterogeneity of the firms belonging to the industry. Of course, such market imperfections as time of delivery, personal contacts, custom, equalizing and other types of discriminatory price systems, influence the localization of industry and distort the findings listed above. These market imperfections tend mainly to yield special theories of location, not the general theory which is required for understanding of underlying trends.[1] At the same time, they suggest that different ethnic and religious groups set up unique factors of consideration that would distort any theoretical presentation of the location of individuals which can be founded on the pure basis of general cost and general utility functions.

[1]And see M. L. Greenhut, "A General Theory of Plant Location," *Metroeconomica*, Vol. VII, 1955, pp. 57–72, for an example of rejection of certain special locational data and subsequent derivation of a generalized theory.

V.

Locations (Plant and Otherwise) in a Privately Capitalistic and a Purely Communistic Economy.

While the maximum-profit location probably represents an approximation to most plant locations in the United States, there arises the question whether the economist ought not evolve explanations of underlying forces behind plant location in any society, thus treating locations in a privately capitalistic society as a special case. Weber was concerned with this question; he sought to explain location forces of any economy. Lösch, on the other hand, did not abstract intentionally from private capitalism; rather, he sought to explain underlying forces in a privately capitalistic economy in which resources and population were evenly distributed. Indeed, Lösch went even one step further by seeking to define the location of things and people in general. In this writer's view, he failed "technically" in his own objective because the ideal system of location that he related is irrelevant to capitalism, even as a norm that follows the special assumptions which he hypothesized for deducing his system. Lösch's rational system of location could be brought about only by full direction from the state. In a sense, he therefore offered a theory of location where the profit motive was either stifled or controlled. Without design, he plugged an open gap in Weber's analysis by non-inclusion of the theory of games.[1]

As noted in section II of this paper, Weber divided cost factors into those that would exist in any country and those that would arise only under private capitalism. Discussion of his factor delineations is not only out of order here, but also unnecessary to present purpose. Suffice it to say that his disregard of the area scattering of consumers led him to define improperly the market area of a firm; it ruled out an appraisal

[1]Lösch, like Lewis (see *supra*, n. 38), offered the hexagonic system as a conception of a rational or ideal system of location. But unlike Lewis, he presented it in the context of a treatise on general location theory which included many references to the need for considering cost and demand data in a single model. Though this background might suggest that Lösch would form a general theory of location, his system lacks capitalistic roots and therefore does not attain full generality.

of locational interdependence, which investigation must be had in order to understand fully the spatial distribution of economic units. Indeed, the closest Weber came to this end was in analyzing the location of different stages of production. But here too, the purely competitive focus prevailed.

Oddly enough, Lösch also failed to analyse locational interdependence. His conceptual framework, while built upon area concepts, conceived primarily of the influence of those competitive forces (disregarded by Weber) which cause intra-industry entrepreneurial scattering, and inter-industry concentration. But just as there are cost and market area differences that would exist at alternative locations in a privately capitalistic economy, there are interdependence factors which would lead to some intra-industry concentration, in even such an economy as one characterized by even distribution of resources. The singular existence of the profit motive, itself, causes uncertainty regarding the location policies of competitors. This factor of consideration promotes localization of an industry.

Indeed, even if it were known that invariably all rivals disperse quasi-monopolistically, the existence of, (1) unequal costs at alternative locations, due, if for no other reason than, to advantages of agglomeration, (2) a demand curve of high inelasticity, (3) a rising marginal cost curve, and (4) a very low ratio between the freight rate and the b (price) intercept, would create uncertainty regarding dispersion and promote concentration of an industry. State planning, perfect knowledge, and complete adherence to the means of maximizing total effective demand for economic goods are prerequisites to the intra-industrial type of scatter conceived by Lösch. These hypotheses are foreign to a fully private capitalistic economy where other conditions and values exist.

It is somewhat surprising that Lösch (the major exponent of the need for treating the cost and demand factors as variables) attempted to abstract from cost by the assumption of a homogeneous land surface. Yet, much more significant, he failed to stress the fact that even under this tenuous assumption there would arise unequal costs at alternative locations. Such differentials follow normally from the difference in trading areas

of diverse products; these differences in market areas make manifest the advantages of agglomeration, even intra-industry.

Difference in sizes of trading areas suggests agglomeration. Agglomeration manifests unequal costs. Accordingly, two firms, a few firms, or all firms, seeking to produce a new commodity at the optimum location may not locate at different places in the short-run, and, most vitally, may be found together (in some form of cluster) in the long-run. Their ultimate trading areas thus will not only *not* take the form of a hexagon, but may be identical. Cost differences make *uncertainty* an *inherent* part of the system; they cannot be omitted as non-integral data. Intra-industry entrepreneurial concentration may exist, even under Lösch's assumptions, just as readily as inter-industry concentration. Most significantly, the localization may be so extreme as to eliminate all trace(s) of the hexagon(s).

While Weber abstracted from area concepts, and thereby failed to consider adequately the problem of maximization of total effective demand, and while Lösch abstracted from the influence of forces causing intra-industry locational interdependence, and thereby failed to appraise adequately the privately capitalistic economy, it is believed that a fusion between the two approaches not only is possible but offers insight to underlying location factors. On the other hand, this fusion will be seen to be incapable of providing full comprehension of the whys and wherefores of location in privately capitalistic economies

In regard to the fusion and initial belief, it is patent that a mixed land surface (heterogeneous (Weber) in parts and homogeneous (Lösch) in other parts) will lead to inequalities in costs (Weber—other than those that arise only from agglomerating advantages) and some spatial scattering of consumers (Lösch—open area and full area sectors). To the extent that Lösch's homogeneous plain exists, intra-industrial scattering and inter-industrial concentration will find some expression, and, in fact, possibly[1] complete expression in a non-free controlled economy. The pattern for state planning and maximization of total effective demand in a totalitarian type of state is

[1]This depends upon whether the savings in freight cost exceed the advantages of agglomeration that arise from the differences in trading area of each good.

thus laid bare, provided further that consumer sovereignty is replaced by state preference (*e.g.* substitution of convenience good buying for shopping good kind of selections).

The same type of fusion does not hold, however, for plant locations in a privately capitalistic economy. Failure is due to the disregard by both Weber and Lösch of forces which account for intra-industry locational interdependence. Private capitalism requires a maximum profit type of theory in which uncertainty (and its minimax principle) is a main variable. Only in this light are the underlying location forces of such economic system revealed.[1] This framework, as outlined in section IV, must be used as the cornerstone for a general theory of location in a privately capitalistic economy. It appears capable of explaining practically all business locations as well as practically all personal site selections. Indeed, the essential difference with respect to non-business location would seem to be that the individual's cost factors are formed mainly by the transportation and land rental charges while his demand (utility) factors are spread over familial and religious relationships and like data. Most significantly, it is proposed that the spatial regularities shown in the socio-population studies,[2] the associated regularities exampled in the investigations of business locations,[3] and the theoretical expectations of spatial regularities,[4] suggest that all location is part and parcel of the same underlying principle. This principle finds its root basis in the

[1]That this locational interdependence will extend to non-zero-sum games (*i.e.*, see J. F. Nash, "Non-Cooperative Games," *Annals of Mathematics*, LIV, 1951, pp. 286–295) as well as the traditional zero sum games can be appreciated by comparing respectively the location cases where there exists some elasticity in demand with those where demand is of zero elasticity.

[2]G. K. Zipf, *Human Behavior and the Principles of Least Effort* (Cambridge : Mass : Harvard University Press, 1949). J. Q. Stewart, "Potential of Popu'ation and its Relationship to Marketing," *Theory in Marketing*, ed. by R. Cox and W. Alderson (Chicago : Irwin and Co., 1950, pp. 19–40). H. W. Singer, "The 'Courbes des Populations.' A Parallel to Pareto's Law," *Economic Journal*, vol. XLVI (June, 1936, pp. 254–263.

[3]A. Lösch, *op. cit.*, Part IV. P. Sargent Florence, *op. cit.*

[4]G. Tintner, "Homogeneous Systems in Mathematical Economics," *Econometrica*, vol. 16, No. 4, (1948), pp. 273–294. W. Isard, "A General Location Principle of an Optimum Space Economy," *Econometrica*, vol. 20, No. 3, (1952), pp. 406–430.

cost and demand factors of the business firm and, their counterpart, the scarcity-utility factors which govern the individual's quest for maximum satisfaction. In very much the same way that Pareto's system of general sociology is a reflection of general economic theory, the location of individuals is a reflection of the location of firms. A general theory of plant location in a private capitalistic economy may very well include the person, with the individual's quest for space, public facilities, neighborhood, and neighbor type creating a force of uncertainty reflective of business location. In any event, the theory of games has a spatial content: the theory of capitalistic location.

Melvin L. Greenhut.

Rollins College, Fla.

The Taxation of Overseas Profits[1]

(1) The Report of the Royal Commission.

The First Report of the Royal Commission on the Taxation of Profits and Income (Cmd. 8761, February, 1953) and chapter 24 of their Final Report (Cmd. 9274, June, 1955) deal with the question whether overseas profits should be treated preferentially for tax purposes. This article attempts a critical appraisal of the arguments and recommendations of the Commission.

The Commission introduce the subject with three general observations :

1. It is difficult to define precisely "overseas income." Where rent is received for a piece of land abroad, or interest for a loan to a foreign debtor, the meaning is clear. But it is blurred when income results from business activities conducted abroad. It is then not always possible to separate costs incurred at home from costs incurred abroad, nor receipts that flow into the mother country from those that do not. For the purpose of double taxation agreements a workable separation has been made. Wherever there is a "permanent establishment" abroad, its receipts are treated as overseas profits. Under the present law the income of a foreign *branch* or *agency* without independent control abroad is treated as overseas income of a U.K. resident, whereas the income of a foreign *subsidiary* is treated as overseas income of a non-resident, *if central management and control are located abroad.*

2. Britain taxes all income that arises in the country (no matter to whom) and all income of her residents (no matter where it arises). Hence a company that resides in Britain is taxed by Britain on all its profits.

3. The Commission then point out that one has to pay attention to what other countries are doing in this field, and that one cannot treat the question of the taxation of overseas profits as a purely domestic problem, for three reasons :

[1]Discussions at seminars of the Johns Hopkins University, the University of Toronto, and Princeton University, clarified several points in this paper and contributed many ideas. A discussion led by Professor J. R. Hicks aroused my interest in the subject.

(a) It is international usage that the country in which profits arise has a prior right to take tax, as against the country in which the owner resides. This is acknowledged in many double taxation agreements.

(*b*) Some charges, such as royalties and export taxes, though in effect levies on profits, are not technically taxes on profits and cannot be credited against domestic tax liabilities.

(*c*) Most other countries give some preference to overseas profits earned by their residents. Only Western Germany taxes foreign profits at the same rate as those arising at home. Some countries grant complete exemption (*e.g.*, Australia, South Africa, Switzerland). The U.S.A. allows a percentage deduction of the tax rate in the assessment of "Western Hemisphere" profits, and income from certain U.S. possessions is taxed on a remittance basis only. There is a good deal of pressure to extend this concession to U.S. companies operating in other countries. In Canada external corporate profits are granted certain tax concessions.[1]

The Commission next outline the problem of how overseas profits should be treated for purposes of taxation. They confess that no agreement existed amongst their numbers and that three points of view were expressed : (*a*) that there should be exemption or reduction of tax for overseas profits ; (*b*) that there should be no change, and (*c*) that some limited scheme should be devised. There was, however, general agreement that income from overseas investment (as distinct from profits of overseas business) should be fully taxed when it accrues to a British resident and that overseas profits distributed to British residents should be fully taxed. Thus the area of dispute is

[1]The exemptions granted by other countries can easily be exaggerated. The two main contributors to foreign investment are the U.S.A. and the U.K. The U.S.A. grants only very limited reductions and American foreign corporations raise the same complaints of handicaps abroad as U.K. corporations. Canada's foreign investment is mainly in the U.S.A. and the U.K. Since tax rates in these countries are about the same as or higher than in Canada, exemption amounts to little more than what would be granted under double taxation agreements. Other countries with little foreign investment find it administratively simpler to forego taxes on foreign income.

confined to *undistributed* overseas profits of limited liability companies. This includes profits of overseas subsidiaries and excludes income from portfolio investment.

The Commission then put the case for exemption. During the oral evidence, Mr. Millard (now Sir Millard) Tucker asked Mr. S. P. Chambers (of Imperial Chemicals and Federation of British Industries) : " . . . what I want is the unanswerable argument that British companies trading abroad should be relieved of their liability to tax . . . In fact, that is what everybody is talking about at the moment, increase your exports, and that is an export, the carrying on of business abroad, is it not?"—Mr. S. P. Chambers : "Yes" (Day 4, 1067).

The business world, not surprisingly, subscribe unanimously to the view that overseas profits should be tax exempt. The Commissioners favouring exemption argue :

1. We need overseas earnings to pay for our essential imports. The present system puts British overseas business into a worse position than its rivals, for two reasons : not only are taxes abroad generally lower, but these are often removed or reduced where rival firms from third countries compete with British firms.

2. The damage, at present, is done through reducing the reserves of the British firms, thus prejudicing their competitive position in the future. How this conclusion is reached from the evidence given to the Commission is not, however, altogether clear, particularly as the evidence is contradictory. Some argued that deficient reserves impede expansion and increase risks. Other witnesses complained that the tax is passed on directly and that British products are thus priced out of the markets. Others again argued that the present system prevents investments that would otherwise have occurred.[1]

3. Several factors which are supposed to mitigate the vicious effects of the impact of taxation are said to be not operative in the desired manner.

(*a*) A company can emigrate by transferring management and control abroad. *But* since 1951 this is illegal except with Treasury consent ; and even if this prohibition

[1]These arguments clearly do not apply to subsidiaries with management and control abroad, for they are not taxed on their retained profits.

should be repealed, migration involves loss of trade, talent, experience and skill. This should be avoided.

(*b*) A company may set up a foreign subsidiary with management and control abroad. It would then qualify for tax exemption until profits are remitted home. *But* this will break the coherence of the group, it may be impracticable (no local residents to be entrusted with management and control), it may, as above, involve loss of economic ties, and finally the profits paid by the subsidiary to the parent, even though not distributed, would be taxable.

(*c*) It has been said (and the Minority Report supports this view) that double taxation relief does adequate justice to overseas income. *But* since British taxes are higher than most foreign taxes, the British firm pays more than either the foreign local firm, or the firm from a third country which gives tax preference ; further, double taxation relief does not apply to royalties and export taxes ; finally, since the foreign country which grants tax concessions to attract business cannot benefit from these under present double taxation agreements, it resents this state of affairs and discriminates against British firms.[1]

There are also *political* and *legal* arguments for exemption which I shall not examine. But there are further *economic* arguments which are not discussed by the Commission.

1. It is often said that risks are greater abroad and that exemption should compensate for this. But taxes do not take account of risk differentials at home. It is by no means clear that risks ought to be equalized in this way. Even if it is admitted that overseas risks ought to be reduced, alternative methods such as guarantees, loss offsets against domestic incomes, or postponement of tax liability would be more effective.

[1]It is sometimes said that progressive tax rates are imposed by foreign countries in order to hit British business, because others do not earn so much. This argument contradicts the view that the situation is highly competitive. It could, of course, be that the British firms are earning the same profit rate but are operating on a larger scale. But then their reserves are likely to be more plentiful and less vulnerable to tax assaults.

2. A related argument emphasizes the fact that costs are higher abroad than at home, because utilities, transport, etc., which are available at home, may have to be put up by the foreign companies. Social overheads and external economies of various kinds are absent abroad. But since the extra costs will also affect the firm's competitors in the foreign market, they cannot be alleged as an extra hardship compared with them. From the point of view of our national interest, there may be cases where subsidies are justified because private and social costs diverge, but this is no reason for subsidising firms to carry on business in unprofitable areas. If a specific region can be shown to be a worthy candidate, selective encouragement of investment there would be indicated.

3. Those who subscribe to the Benefit Principle in taxation argue that a firm operating abroad derives fewer benefits from government expenditure than domestic firms. But if any principle guides our tax code it is that of Ability, not Benefit.[1]

The Commission next argue the case against exemption.

1. One has to distinguish between the claims for exemption that were made on grounds of *equity* from those of *economic policy*. In equity, the relevant comparison is not with foreigners, but with domestic firms. Admittedly, in some cases (royalties and export taxes) the foreign firm may be worse off than the domestic firm, but this is not generally true. There is no case on equity grounds. But equity is expendable. Nobody would claim *fiat equitas, pereat mundus*. If therefore a good case can be made out, equity may have to go by the board.

2. The question of economic policy was discussed by the Commission mainly in relation to the balance of payments : would exemption yield higher export earnings or, rather by implication, would it reduce import expenditure ? The effects on the allocation of resources, productivity and real income were not treated explicitly. The Commissioners who oppose exemption argue :

(*a*) If we want to raise export earnings in relation to import outlays, the disadvantage of firms operating overseas is only a small part of the problem. "It is common to all

[1]It is doubtful whether one can apply the principle of Ability to companies.

businesses which produce for export, and it may even affect industries which produce for the home market, if they have to face overseas competition." (Report 650). But to give such large-scale relief would be impossible to administer, and, even if it could be administered, would lead to such a large loss of revenue as to be ruled out.

(*b*) A concession to overseas profits only would encourage overseas investment and on that score worsen, not improve, the balance of payments. Given the value of exports, an increase in foreign investment reduces the foreign exchange available for imports.[1]

(*c*) It may be that, later, greater foreign investment will increase the value of exports. But it is doubtful whether the current sacrifice is justified by the future benefit. This country "is already investing abroad more than it can afford" (Report 653). Net long-term investment overseas (*i.e.* deducting overseas disinvestment and long-term investment from overseas in the U.K.) ran between 1946 and 1953 at an annual rate of £120 million, and between 1951 and 1953 at an annual rate of £180 million. "These . . . amounts were in excess of the entire surplus on current account in the United Kingdom balance of payments during the same periods, leaving nothing over for the reduction of the short-term debt accumulated during the war or for the building up of currency reserves." (Report 653).

Here the criterion for "what the country can afford" is implicitly the capacity to develop a current account surplus. Obviously other criteria could be advanced : the amount of idle resources at home, or comparative living standards, or domestic investment opportunities. It is rather like the assessment of a man's taxable capacity solely in relation to his liquidity position. But this implicit standard pervades the whole discussion of this subject by the Commission.

[1]This ought to be qualified for overseas investment that consists in ploughing back profits earned by firms abroad. They are not quite in the same way a drain on foreign exchange as new capital transfers.

(*d*) If overseas investment occurs, not at the expense of home consumption but of home investment, our *ability* to produce for export is reduced, even though the *demand* for exports may have been raised by the concession. The Commissioners presumably are thinking of the improved and securer trade connections and the success of selling efforts that have to be set against the foregone increase in domestic productive capacity.[1, 2]

The Commissioners who oppose exemption conclude that those activities should be encouraged which increase exports with a minimum of overseas investment, rather than overseas investment itself.

The recommendation of the majority seems more in the nature of a compromise between, than a conclusion from, the arguments presented. They recommend that a special category of overseas trading concern called an "overseas trade corporation" should be recognized. Its profits should be exempted from tax, as long as they are undistributed. The category must be so defined as not to preclude central management and control in Britain.[3]

Three types of corporation are intended to be covered. 1. The simplest case is that of a corporation that derives all its income from overseas operations but is under the central control of a board of directors in Britain. The talent of management and valuable trade connections, which might otherwise be lost, would thereby be retained.

But the firm may, indeed should, purchase supplies, equipment and replacement in Britain. It may transfer surplus funds to Britain either to buy or to invest temporarily. These are not overseas trading operations, but they are essential to con-

[1]It is not clear how the Commissioners would fit ploughed back profits of overseas subsidiaries into the picture. Although ploughing back is foreign investment, it does not directly sacrifice either domestic consumption or domestic investment.

[2]The case against sacrificing domestic investment in favour of foreign investment is strengthened if the argument on pp. 102-103 is valid.

[3]Presumably this corporation would not be allowed to set losses against the profits of the parent company. Overseas operations that bear, at present, the full weight of U.K. taxation, are permitted to offset losses. It might therefore not pay a firm, particularly in the beginning of a new venture, to form such a corporation.

ducting the business. Hence these activities should not disqualify the firm from attaining the favoured status.

2. A different group of companies are those that extract or produce commodities overseas but bring them home for sale. Such companies are often merely extended arms of domestic enterprises, geographically abroad, but economically domestic. The distinction between trading *with* and trading *in* the country may be difficult to draw. Yet, this type of company is nearest to the highly lucrative traditional type of colonial enterprise. The benefits that the parent countries have derived from this type of enterprise go far beyond an increased supply of essential imports and an improvement in the terms of trade. Indirect and secondary benefits accrued to the investing countries in the form of growing markets, new discoveries, experience, skill, etc. Just because these were only in a geographical sense *foreign* investments were they so profitable. Even from the narrow point of view of the balance of payments, this rather than the first type of enterprise would, in my view, be the most promising candidate for concessions.

3. A third group of companies are those that manufacture in Britain but sell abroad. Their inclusion implies acceptance of the principle that overseas investment has no priority over overseas trade. In this category of overseas trading corporation we find the exporting side of a British manufacturer who does not operate abroad at all. Thus not only the export merchant or agent would be admissible, but the exporting branch of the manufacturer. Clearly there are difficulties in deciding what are the real profits of that branch. But the Commission do not consider these difficulties insuperable. The recommendation to exempt this type of company clearly raises a number of questions, such as should exports be encouraged in this manner by tax reductions? Would it not allow too much abuse? Could anyone who is legally entitled in fact form such a corporation?

(2) The Balance of Payments.

In reviewing the evidence presented to the Royal Commission and some of the arguments in its Report, one is struck by the paradox that foreign investment has been advocated

both as a remedy for the dollar shortage (viz. U.S. foreign investment) and as a remedy for the sterling surfeit (viz. U.K. foreign investment). There is, of course, no necessary contradiction. The effects of an act of foreign investment upon the balance of payments depend upon a set of *primary* and a set of *secondary* factors. The primary effects depend upon the current transfer on capital account, and the current and future flow of returned earnings and capital repatriation, in relation to future foreign investment. The secondary effects are more difficult to formulate precisely. They depend upon the foreign country's increased dependence upon the investing country for complementary services and goods (particularly servicing of equipment and replacement), upon income changes abroad and the additional imports demanded, or the imports replaced, upon the extent to which foreign investment changes consumption habits in favour of the investing country's goods, etc. For any period, the flow-out of foreign exchange may exceed or fall short of the flow-back (primary and secondary). It is conceivable that the net flow-back per dollar invested abroad is sufficiently small, and the net flow-back per sterling invested abroad sufficiently large to make the paradox true.

Some people have been concerned lest American foreign investment aggravate the dollar shortage by creating additional transfer difficulties when profits are repatriated. Professor Domar has shown that, on a fairly abstract level, this need not be so, even on quite plausible assumptions.[1] But analogously, it is quite possible that British foreign investment might create a *negative* transfer problem. It may aggravate Britain's foreign exchange shortage.

The most favourable type of activity that would be encouraged by the proposed concession would be one of the type of the American Oil Companies which is in many respects similar to the colonial type of foreign company. Top management would stay at home. Technicians and middle-level managers would be British, would be paid in sterling and would spend a large part of their incomes on British goods. Equipment and supplies would be purchased in Britain. If one could

[1] E. D. Domar, "The Effect of Foreign Investment on the Balance of Payments," *American Economic Review*, December, 1950.

convert the foreign local employees to British consumption habits, all the better.[1]

Britain's balance of payments can thus benefit, both directly from more and cheaper imports with smaller foreign exchange contents, and also indirectly, by making domestic external economies available to the companies and benefiting from the new experiences, discoveries, skill, technology, and from the growth in demand. A good deal of the highly lucrative pre-1929 foreign investment was of this type.[2]

But clearly, not even all American foreign investment has these attractive features. Consider the type of American Company that invests in British manufacturing. In an interesting article in the *District Bank Review* of September, 1955, Mr. John Dunning shows how American subsidiaries have greatly improved Britain's balance of trade. "So it would seem that American manufacturing firms in Britain, by benefiting from the research and manufacturing experience of their parent concerns and often from established overseas sales and distribution outlets, have helped,—and are helping—to further the growth of Britain's newer exports, and to put this country in the forefront of developments in international industry." At the end of 1950 31 per cent. of American direct investment abroad was export-decreasing rather than import-increasing.[3]

Mutatis mutandis, British foreign investment may encourage the exports and import substitutes of foreigners, to the possible detriment of British export receipts.

One has the impression that before 1929 the typical direct investment abroad increased and cheapened the flow of imports. It was offshore domestic investment, whose primary and secondary benefits accrued to the investing country. Since the thirties, however, another type has grown increasingly

[1]The Fawley refinery construction by Esso showed how even the more traditional English and Irish workers took to American clothing. Obviously Savile Row tailoring cannot have a similar appeal to Iranian workers, nor can Scotch whisky compare, in this respect, with Coco-Cola. But British bicycles and Schweppe's stand a better chance of making converts abroad.

[2]Cf. H. W. Singer, "U.S. foreign investment in underdeveloped areas." *American Economic Review*, Proceedings, May, 1950.

[3]C. P. Kindleberger, *The Terms of Trade*, p. 146.

important : the investment that replaces domestic exports by production of similar products abroad.

It would, however, be necessary to investigate more closely whether, and if so, in what manner, foreign investment is liable to reduce British exports. One could argue that much of the so-called "export-reducing" foreign investment is really in the nature of a salvaging operation : exports would have gone down in any case. To establish plants abroad is one way of saving at least some of the ground that is being lost. If this were true, this type of investment would really be "export reduction-reducing," not export-reducing. Since it is induced by the tariffs and trade restrictions of other countries, this line of argument is plausible. Nevertheless, certain instances of type 1 trade corporation proposed by the Commission might well confirm the fears of those concerned with the balance of payments, even if type 2 may be an answer to their hopes.

These fears may now be summarised. Overseas investment has only limited value as a raiser of export receipts.

(*a*) It worsens the balance on capital account.

(*b*) On current account, an encouragement to the retention of profits abroad constitutes an incentive not to remit to Britain.

(*c*) Foreign exports and foreign import substitutes may be stimulated by the foreign investment, with detrimental effects on British export earnings.

(*d*) When an order might have been placed either with the domestic parent or the overseas trade corporation, the parent has an interest to route the order through the overseas trade corporation to reduce tax payment. This also reduces export receipts.

The objection may be raised that, although foreign investment might have some detrimental substitution effects abroad, we ought to rely on the income effect to make up for them. This will raise the total volume of trade sufficiently to make for higher export earnings, even although foreigners will do more for themselves, and will buy more from each other. Such arguments may have been valid in the past, and certainly were valid under less nationalistic government policies. But to-day

most foreign countries plan to make themselves less dependent on foreign supplies. This reduces the hopes placed in larger incomes and trade.

(3) The Effects on National Income.

(*a*) *Real income and the Terms of Trade.*

The discussion of the Royal Commission is preoccupied with the balance of payments. Yet, any discussion of this kind implies certain views on the effects on real income. The tacit rejection of alternative balance of payments cures to do what the proposed measures are intended to do, implies the view that the alternative would impose greater income or welfare losses on the community. The proposal of foreign trade corporations with privileged status implies the view that, if they make a contribution to the solution of the balance of payments problem, the sacrifices which they entail are not unduly heavy. It therefore seems better to bring out explicitly the effects on real income, if only in order to appraise the social costs and social benefits of this type of solution of the balance of payments problem.

The long-run effects on real national income are anybody's guess. At a fairly abstract level, one may distinguish between domestic investment that is import-competing and export-increasing, and similarly between foreign investment that is import (expenditure)—reducing and export-increasing. All these can make a contribution to solving the balance of payments problem, but their real social costs will be different. The choice for a country between these policies is determined (1) by the relation between additional exports, additional imports, and the terms of trade, and (2) by the future prospects of its terms of trade—apart from variation in its own volume of trade.

Those who are pessimistic about the terms of trade, for either or both of the above reasons, will favour domestic import-substitution and foreign investment that saves foreign exchange. Those more optimistic will favour domestic and foreign export-increasing investment. Others, again, may say that this aggregate approach is too crude and that the choice of investment ought to be determined more by types of products and

by regions from which imports come and to which exports go, than by their total values.

If an adverse long-run trend of the terms of trade is anticipated, the encouragement of extraction and production of raw materials and food in any part of the world that is open to trade will benefit Britain. Such an increase will improve Britain's terms of trade [1] or mitigate a deterioration.

Unfortunately, this is not the typical situation envisaged by the witnesses to the Commission. Most of them were concerned with *exports* sold under fierce competition. This suggests that an expansion might be possible only at more adverse terms of trade. In so far as the worsened commodity terms of trade are the result of productivity increases made possible by the tax concessions, real income does not suffer, and may even grow, although the question is still open whether it could not have been raised more if import-substituting domestic investment had been encouraged instead. But in so far as they are the result of merely pushing up exports, they involve a loss which has to be set against the gain from additional trade.

The case is perhaps clearest where the tax reduction is passed on directly in the form of lower prices. The concession then amounts to an export subsidy. Quite apart from a conflict with our international commitments, this would be foolish if, as Mr. Harrod has stressed, we are already charging too little. But even where the tax is not passed on directly, increased sales will push down prices. The situation would be better if the concessions allowed the companies to innovate and establish temporary monopoly positions. Then increased exports might even be accompanied by improved terms of trade. But tax concessions by themselves are too weak to achieve such revolutionary results.

[1] The relevant terms of trade concept here is that of the factorial terms of trade, including services. A fall in export prices caused by a greater improvement in domestic productivity represents a real gain. On the other hand, a rise in import prices caused by a rise in the rate of return on capital invested abroad represents no loss, if the country owns the enterprise producing those imports. *Vice versa*, constant commodity terms of trade should be corrected for a fall in productivity in export industries, and *e.g.* a new tax on profits earned by foreign enterprises. Both worsen the factorial terms of trade including services, though the commodity terms are unchanged.

In the good old days when comparative cost differences were large, and widely accepted as permanent, any concession to foreign trade earnings would have tended to exploit these advantages more fully. In pre-1914 conditions such concessions might have led to an even faster development and cheapening of British imports extracted by colonial companies abroad and would thus have improved Britain's terms of trade. But with the narrowing of comparative cost advantages, and with the growing conviction that these advantages can be manipulated by policy and can be changed in time, the result is much less certain. Tax concessions in such conditions are just as likely to reduce export prices as to lower import prices, just as likely to foster the development of exports and import substitutes in foreign countries, as the sale of cheaper imports to the home country. While the terms of trade may worsen, the higher foreign incomes may not be spent on British goods, either for lack of demand, encouraged by protection abroad, or for lack of supply, as domestic investment opportunities were sacrificed to foreign investment. If, in addition, the long-term trend of the terms of trade is adverse (and if this is not fully reflected in private profit expectations) the effects of encouraging indiscriminately overseas earnings will have deleterious effects on real income. Extreme terms of trade pessimism would cast doubt on the ability of the proposed policies to contribute towards a solution of the balance of payments problem ; milder pessimism would point to the heavy sacrifices in real income that this solution would demand.

(*b*) *Real Income* : *Profits v. Social Benefits.*

An innocent, not versed in the subtleties of economic analysis, may say : "Foreign investment is bound to be worse than domestic investment. If we invest at home, we get the profits and the wages ; if abroad, only the profits." This would be dismissed as a fallacy. Wages, the text book would tell us, are not an additional gain but reflect the product which we forego when we employ workers on the new project.

Yet, the fallacy is useful in pointing to the possibility that private profit rates at home and abroad do not always reflect fully relative social benefits, because wages do not always

reflect foregone alternatives. Whenever domestic wages exceed the marginal productivity of labour, in the sense of the value of output foregone in an alternative employment, profits to not register the full domestic social benefit of the investment. The excess of wages paid over marginal value product ought to be added to profits to arrive at social benefit. If the investment is abroad, this benefit accrues to the foreigner. Profit comparisons therefore tend, in this case, to under-estimate the advantages of domestic investment.

This point is obviously most important when workers are unemployed. The businessman compares profits at home with profits abroad. But the relevant *social* comparison is profits plus wages (plus multiplier effects) at home with profits abroad. But even in conditions of full employment an excess over the marginal productivity of labour in existing employments arises in a growing economy. First, technical progress may set workers free ; their wages do not reflect any production sacrificed. Population increase and an agricultural reserve may also cause wages to exceed foregone production.

Second, the new investment may itself raise the marginal productivity of labour. The resulting rise in real wages again does not reflect foregone alternatives, but should be added to the return on capital to arrive at the full social gain from the new investment. Whenever wages do not reflect foregone alternatives, be it because wages exceed marginal productivity, or be it that the new investment raises the marginal productivity of labour above the value of foregone alternatives, the benefits from domestic investment will exceed the benefits from foreign investment by the difference between wages paid and the value of the product which the workers could have produced alternatively.[1]

We may conclude this section by saying it is clearly desirable, from a national point of view, to prevent managerial and technical skill from emigrating, and to prevent loss of valuable connections. Beyond that, it is not established that an encouragement of overseas profits will improve the balance

[1]This conclusion would have to be qualified where foreign investment absorbs domestic labour, *e.g.*, skilled workers, technicians, etc. go temporarily abroad.

of payments, and, even if it does, it is doubtful whether the sacrifice, imposed by more adverse terms of trade or by divergence of social benefits from private profits, is smaller than it would have been had alternative policies been adopted to improve the balance of payments.

(4) Help to Colonies.

It would be both selfish and short-sighted to conclude the discussion here. Although beyond the terms of reference of the Royal Commission, many would want to consider tax concessions in the light of their assitance to underdeveloped countries, and particularly to the British colonies. Under present double-taxation agreements concessions granted by these countries fail in their purpose. They merely swell the coffers of the British Exchequer. But one must also remember that the same double taxation agreements allow the colonies to raise their tax rates without having to fear deterrent effects. This is an encouragement to impose heavy taxes on British concerns, but it does no harm to, and indeed, may benefit the colony.

If Britain gave up all tax claims, the Colonies would benefit considerably less than is often thought. They would be competing with each other by concessions to attract capital and would sacrifice revenue in the process. The total of British foreign investment is not likely to go up by much, since more powerful deterrents are simultaneously at work, and a large part of any increase will go to rich foreign countries. Thus any given colony will get very little extra capital, and the poorer, smaller, and more deserving it is, the less it is likely to get, while it sacrifices revenue which it can ill afford to do.

Under the present system, companies are free to benefit from the concessions granted by the colonies if they form subsidiaries with local control. This may be considered desirable from the point of view of the colony. But even then the benefits are lost when the profits are remitted to the parent company, and in any case, subsidiaries are not always desirable or possible.

The Interim Report of the Commission recommended that it should be made possible for the British government to enter into agreements with other countries providing for the grant of credit by Britain against British tax liability for such an amount of foreign tax as the other country certifies to have been spared to the taxpayer by virtue of specified tax concessions in that country. This may go some way in the right direction, but would not cover the situation where a country keeps taxes generally low in order to encourage and attract business.[1]

In any case, the question of aid to underdeveloped countries (which may take many forms other than direct private foreign investment) should be treated separately from the question of exemption of overseas profits. Even if a case for tax concession to firms operating in *poor* countries can be made out on grounds of *aiding these countries*, this would hardly constitute a case for favouring firms operating in both *rich and poor* foreign countries on grounds of British *balance of payments* considerations.

Paul Streeten

Balliol College,
Oxford.

[1]Thus the joint Brazil-United States Technical Commission on Economic Development in Brazil rejected any income tax exemption for new enterprises (except for so-called favoured investments) in its *Report* (1949), but recommended a general low level of income tax rates. Pp. 296-297).

Books Received

DANDEKAR, V. M. : *Use of Food Surpluses for Economic Development.* Gokhale Institute of Politics and Economics. Publication No. 33. Rs. 5 or 7s. 6d. or $1.10. 153 pp.

MURANJAN, Dr. S. K. : *Some Reflection on Economic Growth and Progress.* Gokhale Institute of Politics and Economics. R. R. Kale Memorial Lectures, 1956. Rs. 1 or 1s. 6d. or $0.20. 24 pp.

National Planning Association : *Recommendations for the Future by the N.P.A. Special Policy Committee on Technical Co-operation.* 192 pp. $2.50.

THORNE, Alfred P. : *Size, Structure and Growth of the Economy of Jamaica.* Institute of Social and Economic Research, University College of the West Indies. 112 pp.

The Growth of Life Assurance in U.K. since 1880 [1, 2]

I *Introduction*, *pp.* 109-111.

II *Growth of Life Assurance in U. K.*, *pp.* 111-133.

Annual rates of growth of ordinary and industrial assurance.

Relation of premium income to total personal income.

Sum assured per unit of premium income : comparison with U.S.A.

New business, 1946-54.

Composition of sums assured with respect to type of policy.

Post-war development of " schemes " business.

III *Size distribution of total business*, *pp.* 133-140.

Analysis of the degree of concentration in the industry and the changing relative positions of the major companies.

IV *Costs and Size*, *pp.* 140-146.

Economies of scale.

V *Prices and Profits*, *pp.* 146-152.

Inter-office variations in mortality, interest rates and premiums.

Profit rate and size.

[1] The authors are greatly indebted to Mr. A. Duval, F.I.A., the Actuary of the Co-operative Insurance Society, for his comments on an earlier draft of this paper. Thanks are also due to the staff of the Manchester Central Reference Library for their assistance with Government publications, and to the Librarians of the Institute of Actuaries and the Institute of Insurance for placing their libraries at our disposal.

[2] This paper was read before the Manchester Statistical Society on 14th, November 1956.

VI *Investment of Life Assurance Funds, pp.* 152-164.

Investment field.
Distribution of assets.
Changes in investment policy.

VII *Life assurance and Taxation, pp.* 164-169.

Allowances to policy-holders.
Reliefs to life companies.
Finance Act, 1894, and Married Woman's Property Act, 1882,

VIII *Summary, pp.* 170-171.

APPENDIX I : Accounts of Assurance Companies and the determination of surplus.

APPENDIX II : Table I : Ordinary Life Assurance, 1881-1953.

Table II : Industrial Life Assurance, 1914-1953.

I. *Introduction*

The two classes of life assurance are Ordinary and Industrial. The distinction between these two branches rests mainly on the method of collection and frequency of payment of premiums : in the Industrial branch premiums are collected by agents at intervals of less than two months (usually at intervals of one, two or four weeks), while in the Ordinary branch premiums are paid by the policy holders at yearly, half-yearly or quarterly intervals, or sometimes by monthly Banker's Order.

Life assurance business is conducted by two main types of organisation, namely, companies, which themselves sub-divide into proprietary companies with share capital, and mutual companies with no share capital, in which all profits belong to the policy-holders, and collecting friendly societies. Of the companies some transact ordinary life business only, while others combine this with industrial business, and yet others do general insurance business (fire, accident, marine, etc.) in addition to life. The collecting friendly societies (" voluntary associations or partnerships formed for the purpose of insuring benefits to their members," Bunyon : Law of Life Assurance) concentrate in the main upon industrial assurance, but ten participate also in ordinary life assurance. Table I shows the numbers of companies and societies engaged in life assurance in 1954.

TABLE I

NUMBER OF COMPANIES AND SOCIETIES ENGAGED IN LIFE ASSURANCE, 1954

Class of business transacted	Companies	Societies
L	34	—
IL	1	97
L, IL	5	10
L. plus	47	—
L. IL, plus	9	—
	96	107

Source : Summary of Statements of Assurance Business deposited with the Board of Trade during the year ended December 31st, 1955 (Companies) Report of the Industrial Assurance Commissioner for the year 1954. Stone and Cox insurance Tables 1953.

Notes : L = Ordinary life only.
IL = Industrial life only.
L, IL = Ordinary and Industrial life only.
L, plus = Ordinary life and general insurance.
L, IL, plus = Ordinary and industrial life and general insurance.

Of the 96 companies recorded above, 11 were subsidiaries.

The collecting societies were merely classified into the two groups of industrial life only or industrial and ordinary.

The number of companies has remained fairly stable throughout this century, being 85 in 1900 and 101 in 1937, while the number of societies has declined from 183 in 1924 to 107 in 1954. This is in marked contrast to the situation in the United States when over the same half century the number of life insurance companies increased from about 90 to just under 900. " While the total number of business firms in the United States has increased less than threefold over the last half century, life insurance companies have multiplied tenfold. A considerable number have been formed for special investment

or tax purposes, and for other reasons which differ from the customary objectives of furthering the active sale of insurance ; but even after making a liberal allowance for these, life insurance companies as such have multiplied several times as fast as the business firm population of the Nation."[1]

II. *Growth of Life Assurance in U.K.*

Tables II (a) and II (b) present certain crucial figures depicting the growth of U.K. Life Assurance Companies and Societies.

TABLE II(a)

ORDINARY LIFE ASSURANCE (COMPANIES)

	1900	1922	1937	1946	1953
Number of policies (000) ...	2,011	4,426	6,313	7,420	9,068
Total sums assured (£mn.)	641	1,149	2,161	2,855	5,143
Annual premiums (£mn.) ...	23·5	51·5	99·7	134·5	274·7
Fund (£mn.)	246·1	479·5	984	1,313	2,078

Source : Annual Summaries of statements deposited with the Baord of Trade

Notes : (1) *The number of policies includes the number of insurance policies plus the number of annuity contracts.*

(2) *The figures for Sums Assured and the Fund come from the latest valuation return for each company. A company must make a valuation at least once every five years, but some in fact carry out a yearly valuation and others do so at intervals of less than five years. In the 1953 summary 38 per cent. of the valuations related to end 1953 and another 13 per cent to end 1952 the remaining 49 per cent being spread over the years 1951 to 1949*

(3) *Total sums assured = Life policies in force plus amount of annuities.*

(4) *Premium income here includes income from payments on annuities.*

[1] W. A. Berridge, Significance of the Personal Insurance business in the National Economy, page 94, Chapter 3, Volume 6, of Examination of Insurance Companies published by the New York State Insurance Department.

TABLE II(b)

INDUSTRIAL ASSURANCE (COMPANIES AND SOCIETIES)

	1925	1937	1946	1953
Number of policies (000) ...	—	98,686	118,757	121,502
Total sums assured (£mn.)...	1,032	1,534	2,024*	2,882
Annual Premiums (£mn.) ...	45·2	69·2	103·2	143·1
Fund (£mn.)	183·6	409·7	692·5	969·1

* *Estimated.*

Source : Reports of the Industrial Assurance Commissioner.

Before commenting upon the growth displayed in Tables II (a) and II (b) there are three points to be made clear. Firstly Table II (a) relates only to companies established in U.K. and the figures include the results of the overseas operations of these companies. The statistical returns do distinguish for premiums and annuity payments between the income earned inside U.K. and that earned outside, but no such distinction is available for sums assured, life fund, expenses of management, etc. In 1953 premiums out of U.K. represented about 11 per cent of the total premium income shown in Table II (a) : this represented a return to the pre-World War I position when the percentage in the years 1910-13 was just over 10 per cent ; by 1922 this figure was down to 6½ per cent and the pre-1914 figure of 10 per cent was not in fact regained until 1949.

Secondly, ordinary life business is transacted in U.K. by companies established outside U.K. In 1953 there were 20 such companies with a total premium income of £18 million inside U.K., which is 6½ per cent of the premium income of the British based companies and 3½ per cent of the total premium income of these 20 foreign based companies.

Thirdly, some ordinary life business is transacted by a few of the Friendly Societies, who specialise, of course, on industrial assurance : this part of their business does not come into the statistical net of the Board of Trade Returns or the Reports of the Industrial Assurance Commissioner. An analysis of the

Stone and Cox Insurance Tables for 1952 shows 10 Friendly Societies transacting ordinary life business, with a premium of £4·7 millions, which is less than 2 per cent of the total premium income of companies in 1952.

Thus, in view of the impossibility of adequately splitting up the home and foreign business of either British or foreign based companies, we have decided to concentrate upon the overall activities of the British based companies and societies in the fields of ordinary and industrial life assurance. And further, in view of the inadequate statistical information on the ordinary life business of the Friendly Societies and the negligible size of this business, we have throughout excluded this from our figures and analysis.

Turning now to the growth of ordinary life assurance depicted in Table II (a), we see that between 1900 and 1953 the percentage increases in the four categories were :

Number of policies	350 per cent
Total sums assured	700 per cent
Annual premiums	1070 per cent
Life fund	740 per cent

The two striking features of these data are, on the one hand, the rise in the average sum assured per policy from £319 in 1900 to £567 in 1953 and on the other the fall in the sum assured per pound of annual premium. This latter figure moved as follows :—

1900	1922	1937	1946	1953
£27·3	£22·3	£21·7	£21·2	£18·7

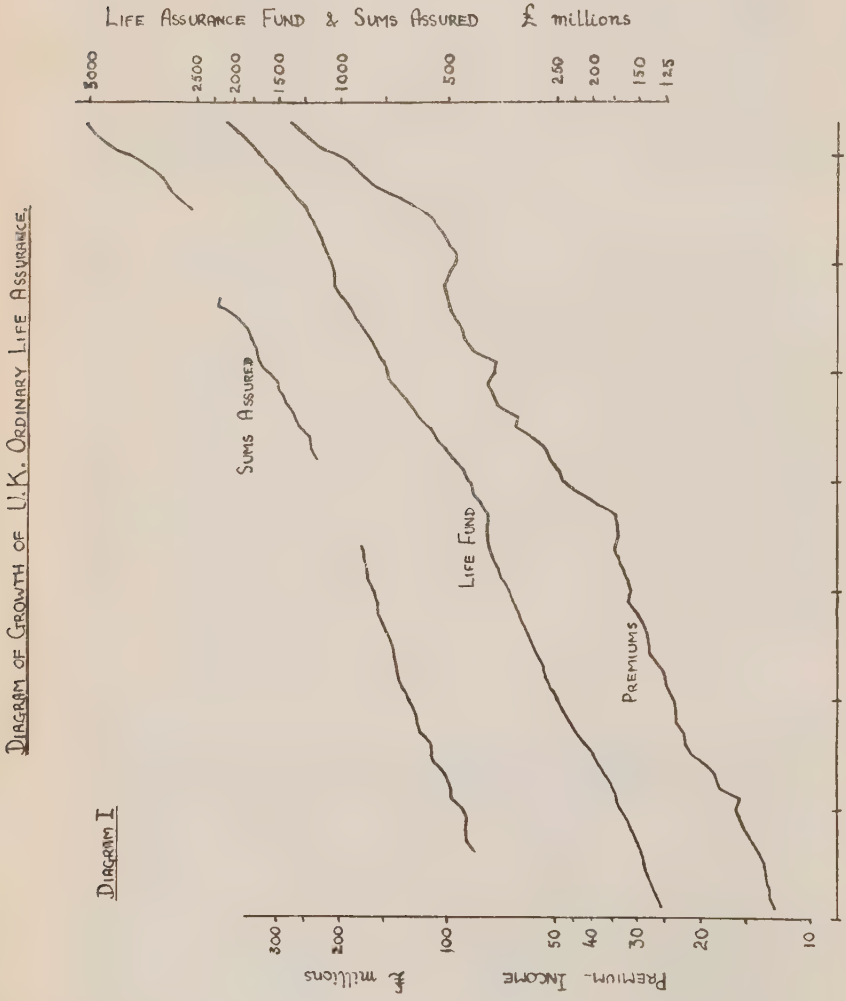
which shows a fall of over 30 per cent. during this period.[1] The

[1] If we exclude the recorded figures for annuities per annum from the total sums assured and the figure for considerations for annuities from the annual premiums in Table II(a) and then calculate sum assured per pound of premium, we obtain the same pattern as before.

1900	1922	1937	1946	1953
£29.3	£23.4	£24.3	£23.1	£19.5

It is impossible, however, to purge the figures completely of annuity business, since *some* offices include as " premiums " in their accounts the amounts received under both life assurance and deferred annuity contracts. But all we are wishing to do here is to obtain a rough measure of the trend in the amount of life assurance provided per pound of premium. The fall in this ratio is due to the changes that have been taking place in the type of assurance provided, and in particular to the shift from whole life to endowment and the growth of ' schemes ' business.

explanation of this fall will be taken up below when we examine the changing composition of life assurance policies in force, and comparisons will also be made with the experience of U.S.A. life offices in this respect.



In Diagram I the growth of ordinary life assurance year by year since 1881 is illustrated and in Table III we present the annual rates of growth of both ordinary and industrial life assurance for selected periods.

TABLE III

ANNUAL RATES OF GROWTH OF ORDINARY AND INDUSTRIAL ASSURANCE

			Period	
		1886–1913	1922–1937	1946–1953
ORDINARY	Sums assured ...	2·8%	4·3%	9·3%
	Premium Income	3·4%	4·4%	10·4%
	Fund	3·6%	4·8%	6·7%
			Period	
			1924–1939	1946–1953
INDUSTRIAL	Companies ...	Premiums	4·0%*	4·5%
		Fund ...	7·7%*	4·6%
	Societies	Premiums	3·8%	4·0%
		Fund ...	6·9%	5·7%
	Combined ...	Premiums	3·7%	4·4%
		Fund ...	6·8%	4·7%

* *1919-1939.*

In the industrial section of Table III the differing rates of growth in premiums and fund in the inter-war period are rather striking : in the cases of both companies and societies the fund was growing at almost double the rate of premium growth. The phenomenon of declining sum assured per pound of premium income is also found in the industrial section, the figures being

1925	1937	1946	1953
£22·8	£22·2	£19·6	£20·1

a decline of 12 per cent. It is also rather remarkable that the sum assured per £ of premium income should now be greater in the industrial than in the ordinary branch, £20.1 as against £18·7. Taking premium income as the basis of comparison, the ordinary section was growing at a slightly faster rate (4·4 per cent) than the industrial (3·7 per cent) in the inter-war period, but since the war the ordinary section has far outstripped the industrial with the remarkable growth rate of 10·4 per cent. A very important contributory factor here is the marked post-war growth of group pension schemes, which are stimulated by existing taxation arrangements (see pages 58 to 62). In the 1946-53 period, the fund in the ordinary life branch has not been growing at anything like the same rate as premium income or sums assured. The disparity between growth in premium and fund depends upon the average age of the policies in force. In the case of any single policy a large part of premium is absorbed by expenses in the early years and only a small contribution is made to the fund : the ratio of premium to fund declines for the policy with the passing of the years. Thus when premium income as a whole experiences a sudden upsurge, as in the years 1946-53, the average rate of growth of premium income will exceed the average rate of growth of the fund in the same period. A relative falling off in premium income will produce the opposite effect, as is illustrated by the experience of the industrial branch between 1924 and 1939.

The overall growth picture is one of a fairly stable number of companies securing a rapidly expanding amount of business : in the ordinary branch premium income grew at an annual rate of 3·4 per cent before the first World War, at 4·4 per cent between the War, and at 10·4 per Cent since the second World War, while in the industrial branch premium income has been growing at about 4 per cent a year, but in this case there is no marked differential between the inter-war and the post-war period.

Another instructive way of looking at the overall growth in premium income is to relate it to the total flow of incomes in the community, as in Table IV.

TABLE IV

RELATION OF PREMIUM INCOME TO TOTAL PERSONAL INCOME

	1938	1946	1947	1948	1949	1950	1951	1952	1953
Ordinary Premium Income as percentage of Total Personal Income	2·01	1·53	1·69	1·73	1·76	1·85	1·92	1·95	2·00
Total premium income as percentage of Total personal income	3·43	2·70	2·88	2·93	2·93	3·01	3·02	3·01	3·04

(Data for Total Personal Income from National Income and Expenditure 1955).

The figures for ordinary premium income in the first line of Table IV show a rise from 1·53 per cent in 1946 to 2·00 per cent in 1953, which is back to the level of 1938 : 1953, however, was a year of full employment and high incomes while in 1938 there was still substantial unemployment. The second row of Table IV gives the figures for the sum of ordinary and industrial premiums as a percentage of personal income. The rise here is less marked, but nonetheless exists. We see, therefore, that the picture of rapid growth in monetary terms displayed in Tables II and III is less pronounced when it is put on a relative basis and contrasted with the growth of total personal money incomes ; but, even here, premium income is seen to be taking a slowly increasing share of the full employment incomes experienced since the war.

It appears that on this basis the U.K. life assurance movement has reached the same relative position as that attained in U.S.A. and Canada. To quote Berridge, " Since 1943, the share of personal income devoted to paying life and annuity premiums has averaged about 3 per cent, remaining fairly constant and

slightly below its 1929 level,"[1] Total premium income of U.S.A. life companies in 1953 was $8·9 billion[2]. Personal income in U.S.A. in 1953 was $286 billion,[3] which gives the premium income of the life companies as 3·11 per cent of total personal income, as compared with 3·04 per cent in U.K. In view of this near equality it is interesting to compare the ratio of sums assured to total personal income in each country. This may be done on two different bases. Firstly, we may take the total sums assured with the U.S.A. and U.K. life companies and compare with the respective personal incomes : these figures will however include sums assured outside the two countries. Total life assurance in force with U.S. companies in 1953 was $318,394 million[4] which represents 111 per cent of 1953 personal income, while the corresponding figure for U.K. is 58 per cent. Alternatively, the total of all life assurance in force in 1953 on lives of U.S. residents, whether issued by U.S. or foreign companies was $305,931 million[5] which is 107 per cent of personal income : the corresponding figure for U.K. was 57 per cent.[6] On either basis the contrast between U.S.A. and U.K. is striking : the sum assured per unit of premium income is roughly twice as great in U.S.A. as in U.K.

It might seem at first sight that this is merely another depressing instance of superior American efficiency, but in the case of insurance this difference can be accounted for by factors which are in no way detrimental to the British industry. An

[1] Berridge, *op. cit.* page 104. Figure 14 on page 105 of Berridge shows that the position in Canada was almost identical with that in U.S.A.

[2] Life Insurance Fact Book, 1954, page 48. This $8·9 billion excludes $1·9 billion for accident and health insurance.

[3] National Income 1954 Edition : Supplement to the Survey of Current business.

[4] Including ordinary, group life, industrial assurance and annuities.

[5] Life Insurance Fact Book 1954.

[6] This was estimated by reducing the sum assured in the ordinary branch with U.K. companies by 4½ per cent to allow for the fact that 11 per cent of the total premium income of U.K. companies in 1953 came from outside U.K. while life premiums going to non-U.K. companies represented 6½ per cent of that premium income (page 4 above) ; to this reduced sum the sum assured in the industrial branch was added and the result expressed as a percentage of personal income.

examination of these factors, however, provides a revealing insight into the differing character of insurance business in the two countries.

Sum assured per unit premium varies between with-profit and without-profit policies. In U.K. in 1953 total sums assured in the ordinary branch amounted to £5,143 million of which 53 per cent represented with-profits policies and 47 per cent without-profits policies. Sums assured in the industrial branch in the same year were £2,882 million. No precise information is available here on the amount of this subject to bonus distribution, but it appears that most offices do make some distribution to industrial policy-holders. " It is still the case in most offices that these allocations to policy-holders are voluntary, because neither the policies nor the prospectuses nor the articles of association of the companies require them to distribute surpluses to industrial branch policy-holders."[1] If we regard all industrial assurance as participating in profits, then the U.K. figures would be 70 per cent participating and 30 per cent non-participating. The figures for U.S.A., kindly supplied to the authors by the Research and Statistics Division of the U.S. Institute of Life Insurance are exactly the same, 70 per cent participating and 30 per cent non-participating, so that this factor does not account for any material part of the difference between U.S. and U.K. figures of sum assured per unit premium. If not all U.K. industrial assurance is regarded as participating in profits, then the U.K. figure for participating policies will be less than 70 per cent and this would be a factor working in the *opposite* direction to the observed difference, since declared sum assured per unit of premium is lower for participating than for non-participating policies.

[1] *The Development of Industrial Assurance in Gt. Britain since* 1928 by K. J Britt in *Proceedings of the Centenary Assembly of the Institute of Actuaries* Volume III, page 13.

Another important cause of variation in sum assured per unit premium is the *type* of policy. This point is illustrated in

TABLE V

SUM ASSURED PER £1 PREMIUM
(Average over 10 largest U.K. Companies, 1953)

Type of Policy	With Profits			Without Profits		
	Whole life	Limited Payment life (20 yrs.)	20 yr. Endowment (at age 30)	Whole life	Limited Payment life (20 yrs.)	20 yr. Endowment (at age 30)
Average Sum ass. at age 30 £	38·7	24·0	18·5	55·5	34·5	22·5

Table V. The sum assured under a whole life policy is twice or two-and-a-half times the sum assured under a 20 year endowment, depending on whether it is a participating or non-participating policy, so we must next see whether there is any difference between U.K. and U.S.A. in the type of life assurance bought. The relevant figures are given in Table VI.

TABLE VI

TYPE OF ORDINARY LIFE ASSURANCE BOUGHT :
U.K. & U.S.A. 1953-54

	Whole Life etc.	Endowment	Other
U.K.	18%	59%	23%
U.S.A.	54%	12%	34%

(U.K. data computed from Board of Trade Returns. U.S.A. data based on figures given in Table VII below).

The comparison in Table VI is based on ordinary branch business only. It is clear that there is a markedly different distribution of life assurance business in the two countries, the outstanding feature being the preponderance of endowment assurance in U.K. (59 per cent) compared with U.S.A. (12 per cent). But this is not the whole story : a factor of 2 or $2\frac{1}{2}$ to 1 for whole life against endowment would not be sufficient, to give an overall difference of 2 to 1 with the pattern of weights shown in Table VI. Therefore the whole life endowment ratio must be more than $2\frac{1}{2}$, or else sums assured per unit premium must be generally higher in U.S.A. than in U.K. To examine the first factor we present in Table VII a more detailed classification of life assurance in force in U.S.A.

In 1954, just under 70 per cent of ordinary business in U.S.A. fell into the straightforward categories of straight life, limited payment life and endowment : the remaining 30 per cent consisted of rapidly growing types of insurance such as term insurance, credit life and family income policies, which show a very high sum assured per unit premium, for in these policies the sum assured may never have to be paid at all, which is not the case in either whole life or endowment. These policies are growing fast in popularity, as evidenced by an 89 per cent increase over the four years in term policies, a 130 per cent increase in credit life and an 85 per cent increase in family income and other combination policies, compared with an increase of 34 per cent in ordinary business as a whole.

TABLE VII

LIFE ASSURANCE IN FORCE IN U.S.A. BY TYPE

Type of Assurance	Dec. 31 1950 Amount $ mn.	Dec. 31 1954 Amount $ mn.	Percentage Increase
ORDINARY			
Straight Life	55,000	70,200	28
Limited Payment Life ...	37,000	43,200	17
Endowment	22,000	25,100	14
Retirement Income with Insurance	9,000	10,600	18
Regular Term	7,000	10,100	89
Decreasing Term*		3,100	
Credit Life	1,000	2,300	130
Family Income and Other Policies	17,000	31,500	85
All others	2,000	4,600	130
Total	150,000	200,700	34
GROUP			
Regular Term	47,000	85,200	81
Credit Life	3,000	7,900	163
All Others	1,000	1,200	20
Total	51,000	94,300	85
INDUSTRIAL			
Straight Life	5,000	6,000	20
Limited Payment Life ...	21,000	24,200	15
Endowment	6,000	5,600	—7
All Others	1,000	2,900	190
Total	33,000	38,700	17
Combined Total	234,000	333,700	43

** Not available separately in 1950 : included in Regular Term. Data supplied by the Institute of Life Insurance, New York.*

Next we see that group policies play a very important and ever increasing part in U.S.A. business, being almost half the size of total ordinary business in 1954, and growing more than twice as fast. This type of policy shows a remarkably high sum assured per unit of premium, as is illustrated in Table VIII.

TABLE VIII

SUM ASSURED PER UNIT PREMIUM. U.S. COMPANIES

Year	Ordinary	Industrial	Group	Total
1938 ...	N.A.	N.A.	N.A.	34·00
1946 ...	33·65	27·78	88·11	35·92
1950 ...	35·74	28·47	94·70	39·66
1953 ...	36·50	26·75	101·73	41·71

(Computed from Life Insurance Fact Book, 1954—A publication of the U.S. Institute of Life Insurance).
N.A.—Not Available.

The group policies are completely dominated (over 98 per cent) by regular term and credit life policies, which accounts for the very high figures shown in Table VIII.

The outstanding feature of the U.S. life insurance field compared with its U.K. counterpart is the marked aversion in the former and the preference in the latter for combining *saving* with *insurance* (endowment policies), and even in the non-endowment field the U.S. companies are now doing relatively much more business in term, credit life and family income policies. Without having detailed information about premium rates in U.S.A. for the various classes of business, it would be unsafe to conclude that the differing composition of business in the two countries is completely responsible for the observed difference in sums assured per unit of premium income, but it does seem probable that it is the major explanatory factor.

A further way of looking at the growth in the monetary value of insurance is to allow for the rise that has taken place in recent years in the general price level. Using the index number of market prices given in the official Blue Book on National Income and Expenditure, 1955, we obtain the results given in Table IX.

TABLE IX

PREMIUM INCOME AND LIFE FUND DEFLATED FOR PRICE CHANGES

Year	Deflated Premium Income (1948 £'s) £mn.	Deflated Life Fund (1948 £'s) £mn.
1938 ...	341	2,860
1946 ...	273	2,305
1947 ...	292	2,288
1948 ...	293	2,260
1949 ...	304	2,348
1950 ...	313	2,394
1951 ...	321	2,349
1952 ...	323	2,341
1953 ...	340	2,477

Note : Premium Income and Fund each refer to the total for Ordinary and Industrial Branches.

We see that premium income in real terms has returned to the pre-war level, while the combined life fund is still about 13 per cent short of the pre-war level ; this is a reflection of the differing rates of growth in the two series, discussed on page 116 above.

Since total sums assured and total premium income are slowly moving aggregates and, therefore, not very sensitive to change it is helpful to look as well at the figures of new business done, presented in Table X.

TABLE X

NEW BUSINESS 1946-1954

Year	New Sums Assured £'000	New Premiums (including single Premiums) £'000	Sum Assured per £1 premium	New Sums Assured as Percentage of Total Personal Income %	New Premiums as Percentage of Total Personal Income %
1946	398,423	20,230	19·7	4·53	0·23
1947	517,257	25,738	20·1	5·49	0·27
1948	511,653	24,697	20·7	5·12	0·25
1949	526,842	24,100	21·9	4·98	0·23
1950	580,988	26,529	21·9	5·27	0·24
1951	686,191	31,635	21·7	5·66	0·26
1952	712,122	30,522	23·3	5·48	0·24
1953	749,892	31,370	23·9	5·46	0·23
1954	872,916	36,084	24·2	6·00	0·25

Source : Board of Trade Returns and National Income Blue Book, 1955.

With the exception of 1948 new sums assured has increased every year, the average rate of increase per year over the whole period being just over 8 per cent. New premiums have also increased, but at the somewhat lower rate of 6 per cent per annum. The result has been an increase in the sum assured per £1 of new premium from £19·7 in 1946 to £24·2 in 1954 : this rise is a consequence of the change in the type of business done in recent years, as will be explained on pages 130 to 133 below.

In times of inflation, most money aggregates will display increases : it is important, therefore, to put them on a relative basis, and this can be done either by expressing the volume of new business as a percentage of total personal income year by year, or by correcting for changes in the value of money. Even when correction is made by either method, *the volume of new business displays a rising trend.* In terms of total personal

income it has risen from 4½ per cent in 1946 to 6 per cent in 1954. Alternatively, correcting new sums assured for changes in the price level, the deflated series has shown *an average annual rate of growth over the post-war period of* 4 *per cent.* New life assurance business is thus more than keeping ahead of inflation, and in real terms shows an annual increase in the business done.

Turning next to the composition of total sums assured, we summarise the main changes in Tables XI and XII.

TABLE XI

SUMS ASSURED—ORDINARY LIFE
Classification by type of policy. £'000

		With Profits	Without Profits	Total
1953	Whole Life	501,224	383,192	884,416
	Endowment	1,995,321	944,960	2,940,281
	Other Classes	115,399	1,013,348	1,128,747
	Total	2,611,944	2,341,500	4,953,444
1937	Whole Life	395,676	207,481	603,157
	Endowment	1,007,898	212,919	1,220,817
	Other Classes	39,375	276,371	315,746
	Total	1,442,949	696,771	2,139,720
1922	Whole Life	370,369	116,805	487,174
	Endowment	419,759	158,318	578,077
	Other Classes	27,485	50,956	78,441
	Total	817,613	326,079	1,143,692
1900	Whole Life	407,385	74,225	481,610
	Endowment	142,393	22,598	164,991
	Other Classes	5,530	23,830	29,360
	Total	555,308	120,653	675,961

Note : These figures relate only to ordinary life insurance : annuities are not included.

The growth in the overall sums assured has been accompanied by marked changes in the composition of the total. Looking first at the classification into policies with and without profits, the percentage composition of the total figure in each of the four years has been :

TABLE XI(a)

TOTAL SUMS ASSURED (ORDINARY)

	With Profits	Without Profits
1900	82%	18%
1922	71%	29%
1937	67%	33%
1953	53%	47%

This movement in the overall figure has been almost exactly reproduced in the whole life class of policy, but not in the case of endowment policies where a cyclical movement emerges.

TABLE XI(b)

ENDOWMENT POLICIES—SUMS ASSURED (ORDINARY)

	With Profits	Without Profits
1900	86%	14%
1922	73%	27%
1937	83%	17%
1953	68%	32%

In the residual category of Other Classes of policy the bulk of the insurance has always been of the without profit type, and in this too the trend has been towards an increase in the percentage without profits.

TABLE XI(c)

OTHER CLASSES—SUMS ASSURED (ORDINARY)

	With Profits	Without Profits
1900	19%	81%
1922	35%	65%
1937	12%	88%
1953	10%	90%

Looking now at the composition of total sums assured by type of policy we find an equally marked change to have taken place.

TABLE XI(d)

TOTAL SUMS ASSURED (ORDINARY) CLASSIFIED BY TYPE OF POLICY*

	Whole Life	Endowment	Other Classes
1900	71%	25%	4%
1922	43%	50%	7%
1937	28%	57%	15%
1953	18%	59%	23%

** If it were possible to fill in a similar table for premium income rather than sums assured, the relative importance of endowment policies in 1953 would be even more marked, because of the low premium rates on Whole Life and Other Classes policies.*

To summarise, the two outstanding features of Table XI are the steady shift towards policies without profits compared with policies with profits, and the enormous drift away from Whole Life policies to Endowments and Other Classes.

A similar, though more recent, change in the composition by type of policy is found in the data for industrial assurance, given in Table XII.

TABLE XII

SUMS ASSURED—INDUSTRIAL ASSURANCE
Classification by type of policy £'000

		Companies	Collecting Societies	Total
1953	Whole Life	1,053,146	404,960	1,458,106
	Endowment	766,866	104,190	871,056
	Other Classes	438,434	34,638	473,072
	Total	2,258,446	543,788	2,802,234
1935	Whole Life	799,047	245,678	1,024,725
	Endowment	232,685	22,196	254,881
	Other Classes	164,673	11,413	176,086
	Total	1,176,405	279,287	1,455,692
1925	Whole Life	559,882	179,044	738,926
	Endowment	216,857	19,948	236,805
	Other Classes	55,693	319	56,012
	Total	832,432	199,311	1,031,743

The percentage distribution of the total sums assured in the industrial branch was :

TABLE XII(a)

TOTAL SUMS ASSURED (Industrial)
Classification by type of policy

	Whole Life	Endowment	Other Classes
1925	72%	23%	5%
1935	70%	18%	12%
1953	52%	31%	17%

Tables XI (d) and XII (a) show why sum assured per unit premium was, in 1953, somewhat lower in the ordinary than in the industrial branches, in spite of the much greater expense ratios in the latter. The explanation lies in the greater preponderance of Whole Life policies in the industrial branch (52 per cent) than in the ordinary branch (18 per cent).

A serious deficiency of the current presentation of life assurance statistics in the Board of Trade Summaries is that the section on Whole Life and Endowment assurances refer to individual policies and everything else is lumped in Other Classes. This last category has grown from 4 per cent of total sums assured in the ordinary branch in 1900 to 23 per cent in 1953 and includes most of the rapidly growing and very important " schemes " business (group life, group endowments with or without pension options, group pensions, etc.). This

TABLE XIII

LIFE PREMIUM INCOME : GROWTH OF " SCHEMES " BUSINESS

	1951	1954	1955 (estimated)
	£	£	£
Group Life	4,991,049	7,460,304	8,200,000
Collective Life	1,082,000	1,307,000	1,375,000
Group Endowment	6,886,616	11,094,617	12,600,000
Individual Policies (under schemes)	17,723,000	27,500,000	31,500,000
Group Pensions	41,822,971	63,229,962	72,000,000
Total " Schemes " Premiums ...	72,505,636	110,591,883	125,675,000
Ordinary Life Assurance Premiums	147,556,840	179,126,227	188,500,000
Aggregate Premiums	220,062,476	289,718,110	314,175,000
Ratio (" Schemes " to Aggregate)	32·9%	38·2%	40·0%

Source : The Policy-Holder, *January 12th, 1956, p. 33.*

deficiency is to some extent remedied by the statistics published by the Life Offices' Association. These illustrate the current position and fast growing importance of group business.

Schemes business has thus increased by over 70 per cent in four years and advanced from less than one-third of total premiums to two-fifths. This rapid growth in large-scale selling of insurance has meant varying fortunes for different types of insurance office. Classifying offices into the four groups of (1) industrial-ordinary, (2) not actively cultivating schemes business, (3) some schemes business, (4) specialists in schemes business, a writer in *The Policy Holder* obtains the results shown in Table XIV.

TABLE XIV

NEW SUMS ASSURED
Classified by Type of Office

	1938	1954	1955
Industrial—Ordinary ...	£49,057,000 (100)	£194,195,000 (396)	£212,324,000 (433)
Not actively cultivating schemes business ...	£51,888,000 (100)	£121,119,000 (233)	£139,653,000 (269)
Some schemes business ...	£54,695,000 (100)	£154,483,000 (282)	£179,241,000 (328)
Schemes specialists ...	£56,944,000 (100)	£350,722,000 (616)	£400,606,000 (703)

Source: The Policy-Holder, *January 12th, 1956, pp. 33-34.*

The figures in parentheses are index numbers with 1938= 100, which show the specialist offices in 1955 doing seven times their 1938 business, compared with index figures of 433, 269 and 328 for the other three groups.

A somewhat different grouping of the companies is also available in *The Policy Holder* : this is a classification of fifty major life offices into three broad groups,

(1) " Composite " offices " whose organisations may be viewed as being primarily concerned with the cultivation of fire and general insurances, but whose ability to provide excellent life assurances has long been demonstrated. These offices have widespread agency connections worked by well-trained and highly qualified inspectors, and the bulk of their life business comes from the United Kingdom."[1]

(2) Industrial-ordinary offices, " mainly concerned with the cultivation of individual policies, largely long term endowment assurances."[1]

(3) Specialist life offices, " devoted to a single purpose, that of cultivating ordinary branch business. The agency system of these offices more closely resembles that of the ' composite ' offices than the industrial ordinary group, being dependent on a net-work of part-time agencies, regularly visited by inspectors who are members of the field staff."[1]

In Table XV is shown the new business (sums assured) of those 50 companies. The position is summed up as follows

TABLE XV

COMPARISON OF NEW BUSINESS PRODUCTION

	1938	1954	1955
Composite Offices ...	£108,106,000 (100)	£453,302,000 (419)	£519,547,000 (481)
Industrial—Ordinary Offices...	£65,183,000 (100)	£230,482,000 (354)	£258,633,000 (397)
Specialist Life Offices ...	£63,744,000 (100)	£186,817,000 (293)	£212,644,000 (333)
Total	£237,033,000 (100)	£870,601,000 (367)	£990,824,000 (418)

Source : The Policy-Holder, *February 16th, 1956. p. 171.*
Figures in parentheses are index-numbers (1938=100).

[1] *The Policy Holder*, February 16th, 1956, page 171.

by *The Policy Holder* :

" Towards this amazing rate of increase the composite offices have contributed the lion's share, and one of the extraordinary features has been the disparity between the experience of the different groups. So far as can be traced, the composite offices have always handled the largest share of the available life business, but it also seems the case that their intimate association with a great many business and industrial firms has put them in a privileged position to advocate pension and similar ' scheme ' business."[1]

This position is summarised in percentage form in Table XVI.

TABLE XVI

PROPORTION OF YEAR'S BUSINESS

	1938	1955
	%	%
Composite Offices	45·6	52·4
Industrial—Ordinary Offices	27·5	26·1
Specialist Life Offices	26·9	21·5

Source : The Policy-Holder, *February 16th, 1956 p.171.*

III. *Size Distribution of Total Business*

The previous section has analysed the growth in the total amount of life assurance by type of policy. We now turn to an examination of the distribution of this total business between the various companies, and we will examine the degree of concentration in the industry at various selected dates and the changing relative positions of the leading companies.

The simplest picture of the degree of concentration in the industry is given by examining the cumulative percentage of the total insurance in force which is done by the top company, the top two companies and so on. This is done in Table XVII.

[1] *op. cit.* page 171.

TABLE XVII

DISTRIBUTION OF ORDINARY LIFE ASSURANCE AMONG THE TOP COMPANIES

Rank of Companies	Cumulative percentage of total assurance in force			
	1913 %	1937 %	1948 %	1953 %
Top Company ...	11·9	12·4	17·3	18·0
Top 2	17·3	18·5	23·0	28·2
Top 5	30·1	33·4	37·7	42·8
Top 10	46·4	48·5	52·6	57·8
Top 20	68·9	67·9	70·6	73·6

As mentioned on page 110, the number of companies has not changed very much, but we see from Table XVII that the degree of concentration has increased over the last forty years. The top company has been the same throughout, namely the Prudential, and its relative share has risen by 50 per cent from 11·9 per cent to 18·0 per cent of the total assurance ; the top two have increased their share from 17·3 per cent to 28·2 per cent and so on down the list.

This increasing degree of concentration is the reverse of the situation in the U.S.A. shown in Table XVIII.

TABLE XVIII

DISTRIBUTION OF ORDINARY LIFE ASSURANCE AMONG THE TOP COMPANIES IN THE U.S.A.

Rank of Companies	Cumulative percentage of ordinary assurance in force	
	1900 %	1953 %
Top 1	16·9	13·5
Top 5	60·1	41·9
Top 10	72·4	55·2

Source : Berridge *op. cit. p. 77*

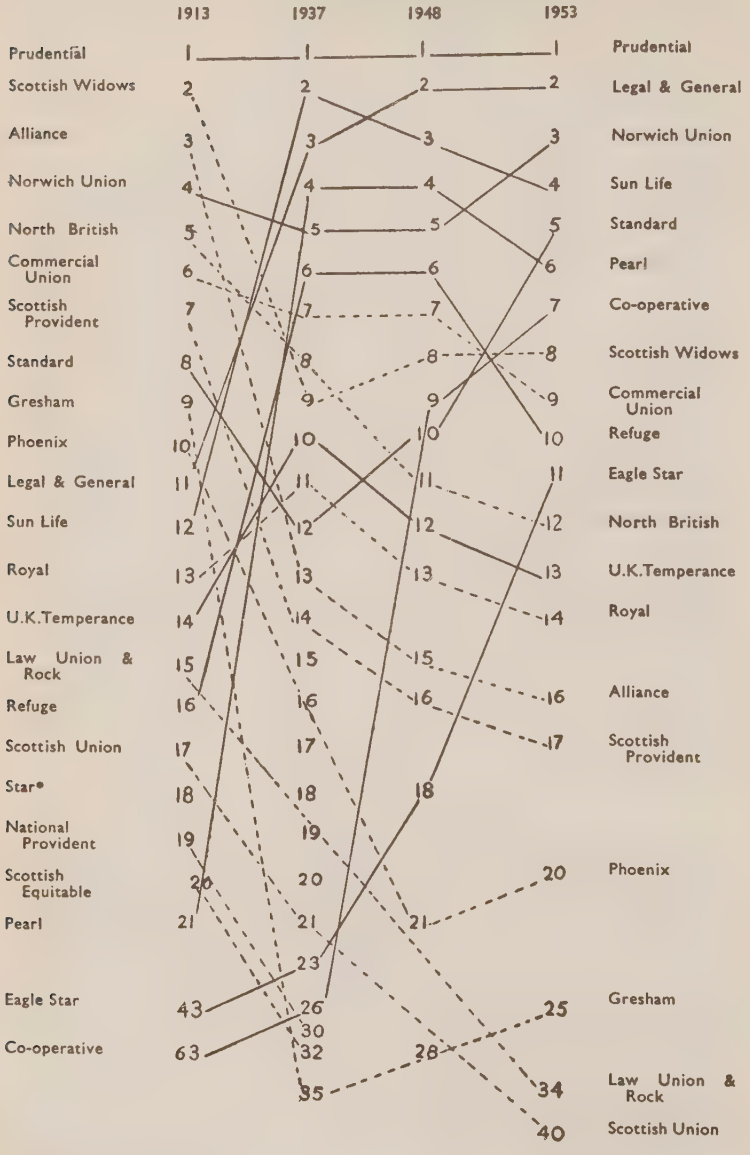
A substantial amount of de-concentration has taken place in the U.S.A., so that the 1953 position, so far as the top five and top ten companies are concerned, is now much the same in each country.

Behind the percentages given in Table XVII several important changes have taken place in the relative position of the leading companies. These changes are illustrated in Table XIX. The ranks shown in this Table are on the basis of Sums Assured, as shown by the latest valuation return for each company. Companies, however, are only obliged by law to carry out a valuation at least once every five years, though some in fact do so at one, two, three or four yearly intervals. This induces some lack of precision in the rankings: for example, if we compare rankings on the basis of sums assured with rankings on the basis of premium income, the same companies constitute the list of the top twelve in each case, but there is one serious discrepancy in the ranking and that is Eagle Star, which in 1953 is sixth on the basis of premium income and eleventh on the basis of sums assured, its latest valuation recorded in the report for 1953 being at December 31st, 1951. Of course, even if valuations were made annually, ranking by sums assured and premium income would not necessarily coincide because of varying emphasis on different types of policies in various companies. Bearing this limitation in mind, we turn to an analysis of the situation pictured in Table XIX.

Companies which have maintained or improved their position in 1953 compared with 1913 are shown with a full line and those which have lost ground relatively are shown with a dotted line. Of the ten leaders in 1913 only three did not lose ground: the Prudential remained in majestic command throughout, the Norwich Union dropped one place and then gained two, while the Standard Life dropped four and then gained two and five places in succession to record a net gain of three places. The relative decline of the other leaders is matched (or caused) by the strong upsurge of half-a-dozen companies,

TABLE XIX

RANKS OF THE LEADING COMPANIES (ON BASIS OF SUMS ASSURED).



*Merged with Eagle Star, 1917

namely the Legal & General, Sun Life, Refuge, Pearl, Eagle Star and Co-operative. Table XX shows the rates of growth for these companies over the period.

TABLE XX

RATES OF GROWTH OF INDIVIDUAL COMPANIES, WHICH HAVE IMPROVED THEIR POSITION IN 1953 COMPARED WITH 1913

Company	Sums Assured		Annual Rate of Growth %
	1913 £'000	1953 £'000	
Legal & General	23,260	523,800	8·1
Sun Life	23,230	263,100	6·3
Norwich Union	35,390	270,600	5·2
Standard	29,480	221,000	5·2
U.K. Temperance	21,730	87,270	3·5
Refuge	19,430	135,900	5·0
Pearl	13,460	199,500	7·0
Eagle Star	5,397	117,400	8·1
Co-operative	724·5	154,300	14·4
Total for all Companies ...	858,160	5,143,000	4·6

Table XX emphasises the remarkable growth of the Co-operative, with a rate of 14·4 per cent per annum over the 40 years, compared with an average annual growth of 4·6 per cent for all companies: while the Legal & General and the Eagle Star had rates of growth almost double the average rate for all companies. The relative positions of the leaders, however, appear to be changing fairly continually. If we look at the shorter period of 1948-53, the annual rates of growth are given in Table XXI.

TABLE XXI

ANNUAL RATES OF GROWTH OF SOME LEADING COMPANIES
1948-53

Company	Rate of Growth %
Legal & General	22·8
Sun Life	8·4
Norwich Union	13·6
Standard Life	23·1
U.K. Temperance	5·1
Refuge...	3·5
Pearl	4·9
Eagle Star	18·7
Co-operative	10·0
All Companies	9·6

The Co-operative now appears to be growing at just a little more than the average rate, while the Legal & General, Standard Life, Eagle Star and the Norwich Union are shooting ahead due to the large amount of " schemes " business which they do. Many other smaller companies have also achieved very high rates of growth in recent years.

It would appear that in most cases the growing companies have grown simply by expansion of business rather than by absorption of or association with other companies. The major exception to this remark is the Eagle Star. Out of ten amalgamations since 1913 by the companies listed in Table XIX, five were made by the Eagle Star, two by the Commercial Union, and one each by the Prudential, Sun Life and Alliance. Twenty-two companies were taken into association with these major life companies since 1913, six with the Commercial Union, four each with Eagle Star and Phoenix, two each with Royal, Scottish Union and Sun Life, and one each with Alliance

and North British. These figures actually include non-life companies, and of the companies so far mentioned only Eagle Star and Sun Life have improved their ranking between 1913 and 1953, while the Legal & General, Norwich Union, Standard, Pearl, Co-operative and Refuge have all improved their rankings without assistance from amalgamations or association.

When we examine the distribution of the industrial assurance business between the participating companies we find that although the total number of companies involved is small (14 in 1953) and has remained fairly steady, a considerable degree of deconcentration has taken place, the reverse of the situation in the ordinary branch.

TABLE XXII

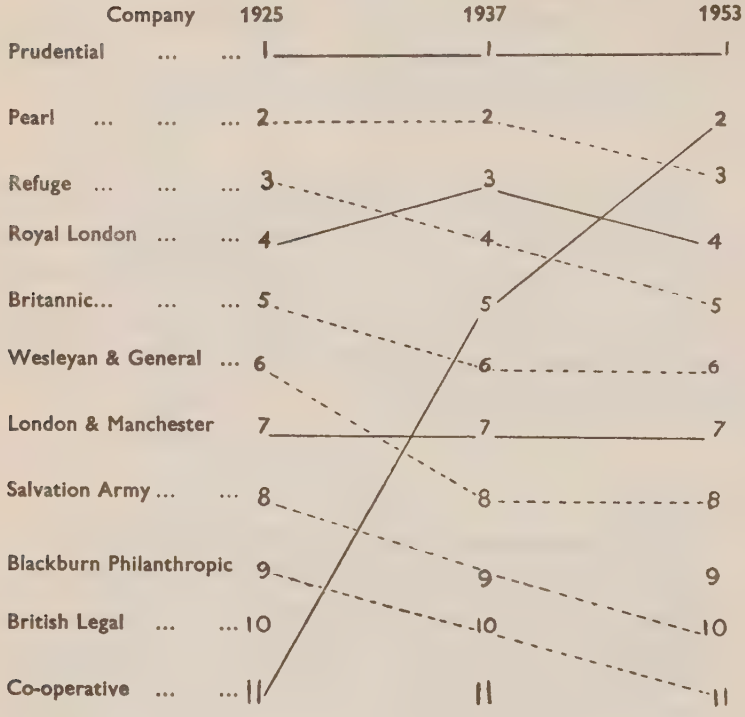
DISTRIBUTION OF INDUSTRIAL LIFE ASSURANCE BUSINESS AMONG THE TOP COMPANIES

Rank of Companies	Cumulative percentage of total assurance in force			
	1900	1925	1937	1953
Top 1	58·6	53·5	46·4	42·8
Top 2	70·2	66·4	59·6	57·0
Top 5	91·8	87·8	85·8	85·3
Top 10	99·5	98·3	99·1	99·8

The top company has again been the Prudential throughout, while the others have been interchanging places behind, as is shown in Table XXIII.

TABLE XXIII

RANKS OF INDUSTRIAL ASSURANCE COMPANIES
(On basis of Sums Assured)


Company
1925
1937
1953
Prudential 1 1 1
Pearl 2 2 2
Refuge 3 3 3
Royal London 4 4 4
Britannic... ... ... 5 5 5
Wesleyan & General ... 6 6 6
London & Manchester 7 7 7
Salvation Army 8 8 8
Blackburn Philanthropic 9 9 9
British Legal 10 10 10
Co-operative 11 11 11

The only company that improved its relative position between 1925 and 1953 was the Co-operative, which moved in spectacular fashion from 0·6 per cent of the total business in 1925 to 14·2 per cent in 1953.

IV. *Costs and Size*

One of the questions in which economists are naturally interested is whether any relationship exists between the size of an organisation and the level of its costs. Are large organisations more or less efficient than small ones, for example,

in the sense of having lower or higher costs ? The following analysis is confined to ordinary life business, where costs consist of all management expenses and commission, the sum of which is expressed as a percentage of premium income to give an expense ratio. We take this ratio as defined and calculated in the Stone and Cox Insurance Tables.[1]

The Stone and Cox expense ratio is calculated in such a way as to make a reasonably fair comparison between companies which transact varying amounts of single premium business.

In Diagram II we show the expense ratio plotted against premium income for 1952. There appears a fairly clear negative relation between the two, the expense ratio declining, on the average, with increasing premium income. Six companies appear on the graph with very low expense ratios ; they pay no commission. The correlation coefficient between expense ratio and the logarithm of total annual premiums is —0·3072 which, for 61 observations, is significantly different from zero at the 5 per cent level. A similar calculation for 1938 yields a correlation coefficient of —0·3930.

[1] " In calculating the expense ratio we have deducted the single premiums from the total premium income of the year, and 5 per cent of the amount of the single premiums from the total amount of commission and expenses shown in the reports." Stone and Cox Insurance Tables (Ordinary Branch) 1953, page 6.

DIAGRAM II: EXPENSE RATIO AND SIZE OF COMPANY (1952)
EXPENSE RATIO (%)
PREMIUMS (£'000)
50000
25,000
5,000
2,500
500
250
50

Classifying the 1952 results into three groups on the basis of the premium income, we obtain the results shown in Table XXIV.

TABLE XXIV

EXPENSE RATIO AND SIZE OF COMPANY (1952)

Total Annual Premiums	£50,000 & under £500,000	£500,000 and under £5,000,000	£5,000,000 and over	All Companies
Number of Companies ...	9	41	9	59
Average Expense Ratio % ...	19·3	15·9	12·0	15·81

The data of Table XXIV and a " significant " correlation coefficient do not prove the existence of lower costs as a *consequence* of larger size. There are two important and relevant factors, which require examination. Firstly, an important influence on costs in any year is the amount of new business done in that year : this factor can be allowed for by working out for each company the premium income from new business as a percentage of total premium income and bringing this additional variable into the correlation analysis. Secondly, most of the " scheme " business, which has developed to such an extent in recent years is done at very low expense rates compared with individual policy business. Thus it is possible that large companies may have a low overall expense ratio merely because of a large amount of " schemes " business. A factor, however, which suggests that this is not the whole explanation for low expense ratios in the larger companies is the significant negative correlation for 1938, a year when schemes business had not developed to anything like its post-war importance. A possible method of dealing with this second factor is to divide the companies into fairly homogeneous groups and to carry out the analysis separately for each group. We have taken the grouping given in *The Policy Holder* of composite offices, industrial-ordinary offices, and specialist life offices and for each group the three variables analysed are :

X_1 = Expense ratio per cent.
X_2 = New premium as percentage of total premiums.
X_3 = Log of total premiums.

The results of the analysis are given in Table XXV.

TABLE XXV

EXPENSE RATIO AND SIZE OF COMPANY
Three Groups (1952)

Group	No. of Companies	Correlation Coefficients				
		r_{12}	r_{13}	r_{23}	$r_{12\cdot3}$	$r_{13\cdot2}$
Composite	25	0·2184	–0·5534†	–0·2945	0·0696	–0·5245†
Industrial ordinary	7	–0·1035	0·0674	–0·3068	–0·0872	0·0376
Specialist life ...	16	0·5617*	–0·6192†	–0·1843	0·5800*	–0·6342†

* *Significant at 5% level*
† „ „ 1% „

We can leave out of account the seven industrial-ordinary companies in Table XXV ; the number is very small and all the correlation coefficients are insignificant. For both the composite and specialist life companies we see that there is a positive correlation (r_{12}) between the expense ratio and the proportion of new business to total business, which agrees with expectation ; there is a strong negative correlation (r_{13}) between expense ratio and size of company, and there is a suggestion (r_{23}) that, on the whole, new business is a somewhat smaller proportion of total business in the large companies than in the small. The crucial question is then whether the negative correlation (r_{13}) between expense ratio and size is evidence of real economies of scale, or whether it is simply due to the smaller relative amount of expensive new business in the larger companies. The answer is given in terms of the partial correlations ($r_{13\cdot2}$) which show the net relation between expenses and size when the effects of the varying amount of new business has been eliminated. Both these coefficients are large and negative,

—0·5245 for the composite companies and —0·6342 for the specialist life offices, which suggests the existence of real economies of scale in both groups.

An even better grouping of companies than that provided in Table XXV would be one which classified companies solely according to their degree of participation in "schemes" business. We were fortunate in obtaining a classification of 35 major companies into those :

(1) Transacting no schemes business.
(2) Transacting some schemes business, and
(3) Specialising in schemes business.[1]

For each of these groups a similar analysis was carried out to the one reported in Table XXV above. The results are given in Table XXVI.

TABLE XXVI

EXPENSE RATIO AND SIZE OF COMPANY (1952)

(Companies grouped according to their degree of participation in schemes business)

Group	No. of Companies	Correlation Coefficients				
		r_{12}	r_{13}	r_{23}	$r_{12\cdot3}$	$r_{13\cdot2}$
No schemes business ...	13	0·3157	–0·4680	0·1622	0·4491	–0·5546*
Some schemes business ...	11	0·2708	–0·3622	0·2176	0·3843	–0·4482
Specialists in schemes business ...	11	0·2661	–0·7518†	–·6435*	–0·4313	–0·7868†

X_1 = *Expense ratio* (%)
X_2 = *New premiums as percentage of total premiums*
X_3 = *Log. of total premiums*
* *Significant at 5% level*
† " " *1%* "

The same general pattern emerges in each of the three groups. Expense ratio tends to fall with size (r_{13} negative) and to rise with the proportion of new business to total business. When we consider the net relation between expenses and size, with the proportion of new business held constant, all three

[1] We acknowledge our indebtedness to the Editor of *The Policy Holder* and his contributor Lifeman for this classification.

partial correlation coefficients are negative ($r_{13 \cdot 2}$), the levels of significance being 5 per cent, 10 per cent and 1 per cent respectively, and in each case the partial coefficient has a greater numerical value than the simple coefficient (r_{13}). These results support those found in Table XXV and suggest that, when full allowance has been made for other relevant factors, there still remains a significant relation between size and expenses, the expense ratio declining on the average with increasing size of company.

V. *Prices and Profits*

After costs it is natural to look at prices and profits. The outstanding feature about prices is the very large number of different prices according to the particular type of insurance bought. Price (i.e., premium) will vary according to

(1) Type of policy (e.g., whole life, limited payment life, endowment, etc.).

(2) Age of insured, term of insurance, and

(3) Whether the policy is on a with or without profits basis, premium rates being lower for the latter.

But even when these three factors are constant (as for example, with a 20 year endowment without profits taken out at age 30) there is still a remarkable range of premiums charged by different companies, because of the differing importance they attach to the determining factors of mortality, interest and expenses. The basic principle for the determination of *net* premiums is " that the present value of the premiums must always be equivalent, at the outset, to the present value of the benefits to be provided by the insurance."[1] In the determination of these present values the two factors of expected mortality and interest loom large. Most companies have in recent years used the A 1924-29 mortality table as the basis of their calculations, but they are, of course, at liberty to vary premiums from the level indicated by this table to allow for variations in their expected mortality from the average. Variations in mortality experience between companies is fairly substantial, as is shown by Table XXVII.

[1] H. E. Raynes, *Principles of British Insurance*, page 85.

TABLE XXVII

CONTRIBUTING OFFICES ANALYSED ACCORDING TO STANDARDISED DEVIATION OF THE PERCENTAGE OF ACTUAL TO EXPECTED DEATHS WITHIN AGE GROUP 45-75 FROM THE CORRESPONDING PERCENTAGE FOR ALL OFFICES COMBINED

Standardised Deviation	Medical No. of Offices	Non-Medical No. of Offices
	Positive Deviations	
7·0—8·0	—	1
6·0—7·0	—	—
5·0—6·0	1	—
4·0—5·0	1	—
3·0—4·0	—	1
2·0—3·0	4	1
1·0—2·0	10	6
0·0—1·0	9	2
	Negative Deviations	
0·0—1·0	10	3
1·0—2·0	8	3
2·0—3·0	5	2
3·0—4·0	4	2
4·0—5·0	2	3
5·0—6·0	—	—
6·0—7·0	—	1
Total	54	25

Source : Journal of the Institute of Actuaries, *Vol. 82, Part 1, p. 36.*

This Table is based on the mortality experience during 1949-52 of the offices contributing data to the Joint Mortality Investigation Committee of the Institute and Faculty of Actuaries. For each office the actual deaths (a) are expressed as a percentage of the number of expected deaths by the A

1924-29 mortality table (E) ; the deviation of this figure from the corresponding percentage for all offices combined (Medical and non-Medical business being treated separately) was then computed and divided by its standard error to allow for variations in the size of the office ; the resultant figure is the standardised percentage given in the table. In the Medical section, seventeen of the fifty-four offices have deviations greater than twice their standard errors, while in the non-Medical section eleven of the twenty-five offices differ from the mean by more than twice the standard errors. In the words of the Committee, " All these observations tend to support the view that inter-office variations are inherent rather than due to chance fluctuation."[1] These variations in mortality depend on such factors as the occupations, class and area for which a given office caters and there is undoubtedly some variation in premium rates to allow for variations in mortality experience.

Another important factor in causing variations in premium rates is the rate of interest that the company can earn on its funds. Some indication of the variation between companies in

TABLE XXVIII

INTEREST RATE (GROSS) EARNED BY LIFE OFFICES
(1952 and earlier)

Rate	No. of Companies
£ s. d.	
3 0 0 to 3 10 0	1
3 10 0 to 4 0 0	8
4 0 0 to 4 10 0	34
4 10 0 to 5 0 0	16
5 0 0 to 5 10 0	5
5 10 0 to 6 0 0	1
	65

Stone & Cox Insurance Tables, 1953

[1] *op. cit.* page 35.

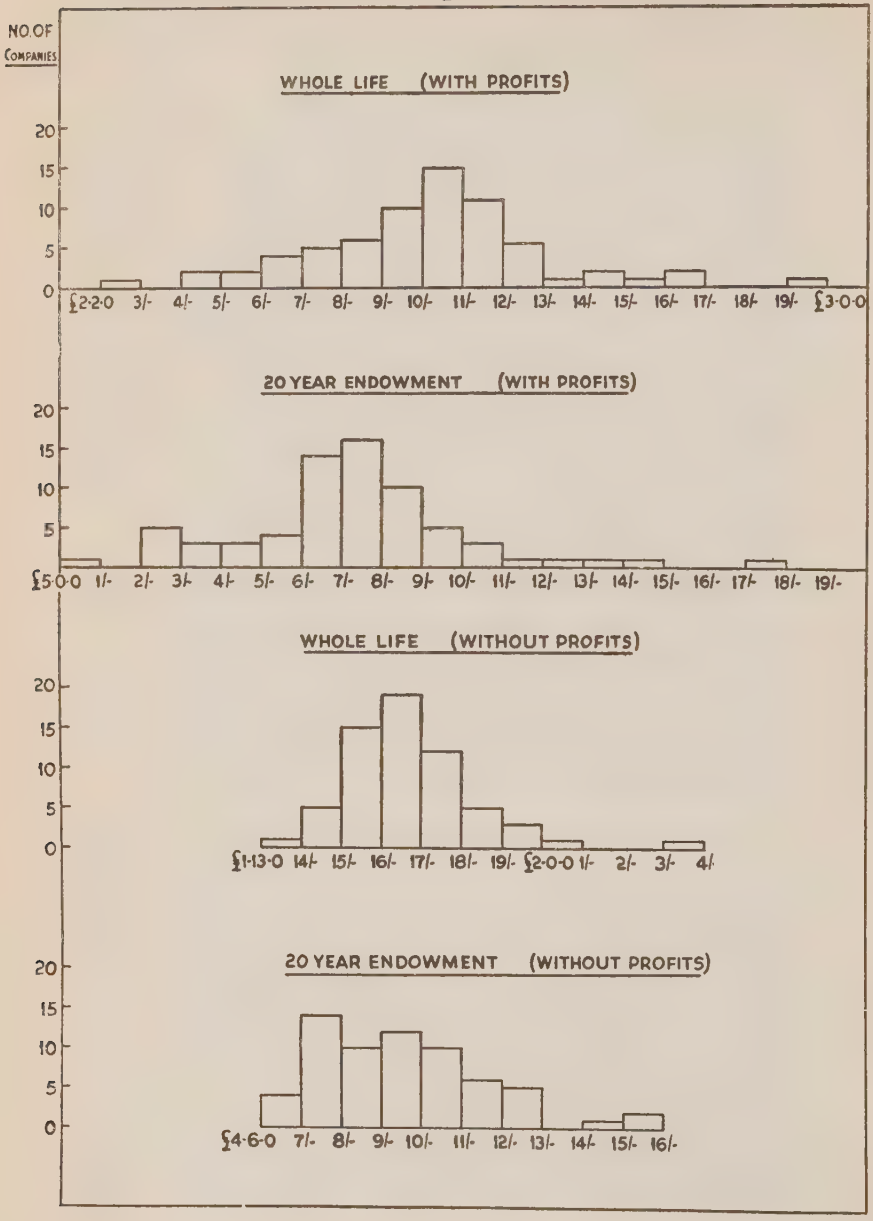
this respect is given in Table XXVIII, though it must be remembered that the variation in the interest rates assumed by companies in the calculation of premiums may not be as great as that in the rates earned.

A third important source of variation, however, is in the loadings added to net premiums in order to obtain the premiums actually charged. The difference between gross and net premiums is mainly reserved for covering anticipated expenses and there is usually a reserve margin for contingencies. Even allowing for these factors the variation in premiums from office to office appears greater than might have been expected. These variations are shown in Diagram III, where all rates are quoted for policies at age 30 next birthday. For whole life without profits the rate of premium per £100 assured ranges from £1 13s. 0d. to £2 4s. 0d., but there is a fairly heavy concentration of more than two-thirds of the observations in the range £1 15s. 0d. to £1 19s. 0d. For a 20-year endowment without profits the premium ranges from £4 6s. 0d. to £4 16s. 0d.

When we turn to policies with profits there is an additional reason for variation in premiums ; this is the loading added to earn the profit for subsequent distribution. The range is in each case practically double the range for the without profit policies : for whole life with-profits it is from £2 2s. 0d. to £3 0s. 0d., and for a 20-year endowment £5 0s.0d. to £5 18s. 0d., although there is still an observable central peak in each distribution.

Companies differ enormously in the proportion of with-profits business which they do. This is a complicating factor when one comes to look at profits, since premiums will have been adjusted to varying extents to earn the surpluses required for distribution : some companies may well charge higher premiums in order to distribute greater surpluses. Given the level of premiums charged, additional factors which may influence profits are differences between actual mortality, interest and expenses experience and the assumption made about these in fixing the premiums. Profits, however, are even more arbitrary in life assurance than in most commercial activities, since they depend upon two sets of estimates, namely estimates of future

DIAGRAM III. PREMIUM RATES PER £100 ASSURED (1952) AGE 30.



liabilities and estimates of future income, and companies may vary in the degree of caution which they exercise in making these estimates. With these warnings in mind we may look at the figures in Table XXIX.

TABLE XXIX

PROFIT RATES IN VARIOUS COMPANIES

Company	Total Sums Assured (Dec. 1953) £'000	Percentage of With-Profits Business	Profit Rate %
(1)	(2)	(3)	(4)
Prudential	908,944	74·4	21·2
Legal & General	456,423	6·4	2·8
Norwich Union	268,817	40·3	14·4
Sun Life	260,240	33·2	9·6
Pearl	199,090	73·0	30·3
Standard	182,150	65·1	9·4
Co-operative	153,817	91·6	23·2
Refuge...	135,784	90·0	49·3
Scottish Widow's	135,024	64·7	29·6
Commercial Union	129,678	29·7	15·6
Eagle Star	103,358	7·3	3·2
North British	87,868	57·5	21·0
United Kingdom	87,132	63·5	32·3
Royal	84,736	48·8	37·2
Britannic	78,585	87·3	33·8
Scottish Provident	74,039	59·1	23·4
Alliance	73,927	43·2	32·1
Royal London	70,873	86·5	21·6
Yorkshire	62,913	27·6	17·6
Phoenix	60,613	32·1	16·9
Royal Exchange	59,872	44·9	27·4
Liverpool & London & Globe	59,714	49·5	32·8
Equity & Law...	54,274	36·7	16·2
London & Manchester ...	54,056	88·7	42·5
Gresham	53,910	53·7	10·4
Scottish Amicable	51,940	61·5	15·3
Atlas	51,929	26·1	37·0
Friends' Provident	50,686	24·2	19·2
Clerical, Medical & General	47,730	48·7	25·6
London Assurance	45,809	31·5	16·9

Notes: Columns (1), (2) *and* (3) *from* The Policy Holder, *December 1st, 1955, p.1441. Profit rate in column* (4) *defined as Surplus ÷ Premiums, and computed from the Stone & Cox Insurance Tables for 1953.*

The companies in Table XXIX are the thirty largest. While sums assured and percentage of with-profits business relate to December 1953, the profit rate in Column (4) is computed from the latest valuation returns given in the Stone

and Cox Insurance Tables for 1953. The period covered by the valuation may vary from one company to another from one up to five years. Taking the log of sums assured as our measure of size and carrying out the usual correlation analysis, we find a simple correlation of 0·5679 between profit rate and the percentage of with-profits business, which accords with expectation. The partial correlation between size and profit rate with the percentage of with-profits business held constant is —0·3150, which just falls short of the 5 per cent significance level. The data suggest a tendency for profit rate to decline with size of company, but the relationship is not statistically very strong and it would be desirable to carry out a similar analysis for other years to see if this pattern persisted.

VI. *The Investment of Life Assurance Funds*

Total assets of life assurance offices established within Gt. Britain have now grown to £4,093,300,103 as shown by returns for the accounting year ended December 31st, 1954. In 1880 total assets were £159,937,000. Examination of the ratio scale graph of the 74 years between these dates shows the steady percentage increase over the whole period. The art of life assurance business is allied to the art of financial investment. How the life assurance companies have invested these large funds over the years is shown on the accompanying charts and tables.

Before an examination is made of the securities in which life companies have invested their funds, it is perhaps useful to consider the relative size of other institutional investors and the aggregate funds and securities available for investment. Unlike United States and Canadian companies, there are no investment rules for British life companies to consider other than the Prudent Man rule. The availability of securities, the requirements of maturity dates, the exigencies of taxation, the

actions of competing institutions, must all have a profound effect on the pattern and trend of life assurance companies investment portfolios.[1]

Table XXX shows the assets of large institutional investors.

TABLE XXX

ASSETS OF SOME INSTITUTIONAL INVESTORS

	1938 £mn.	% Total	1952 £mn.	% Total
Clearing Banks	2,536	50·3	6,667	55·8
Ordinary Life and Industrial Companies ... within Gt. Britain	1,654	32·8	3,528*	29·5
Building Societies	758	15·0	1,478	12·4
Assurance Companies excluding Life	90	1·8	281*	2·4
	5,038	99·9	11,954	100·1

* *Source : Annual Abstract of Statistics*

A much smaller amount of the above is invested in Stock Exchange securities. Table XXXI shows the amount invested in Stock Exchange securities and this amount expressed as a percentage of the total assets shown in Table XXX.

[1] The institutional investors who offer the major competition to insurance companies in the investment markets are pension schemes trusts which are not administered by the insurance companies. Appendix V of the Report of the Committee on the Economic and Financial Problems of the Provision for Old Age (Cmd. 9333) estimates the accumulated funds at £1,228 million.

TABLE XXXI

APPROXIMATE ASSETS OF SOME INSTITUTIONAL INVESTORS IN STOCK EXCHANGE SECURITIES

£million

	1938	% Total Assets	1952	% Total Assets
Clearing Banks	637	25·1	1983	29·7
Assurance Companies (Life and other Companies with in Gt. Britain)	1,298	74·4	2,687	70·5
Building Societies	53	7·0	158	10·7

Total Value of Stock Exchange securities 31st March, 1953, amounted to £27,000 million and this can be apportioned as follows :—

British Government Securities...	£15,000 million
Commercial and Industrial ...	£5,000 million
Others	£7,000 million
	£27,000 million

From these figures it is clear that insurance companies hold approximately 10 per cent of Stock Exchange securities. The figures show the enormous preponderance of British Government securities. In March, 1938, the market value of British Funds was £6,859 million representing 39 per cent of securities quoted in the Stock Exchange Official List, whereas in March, 1953, British Funds were absorbing well over 55 per cent of the Stock Exchange quoted funds.

We can see from the tables the possibility of dependence of the institutional investor on British Funds in the post-war period.

What has actually happened to the distribution of the assets of British Life companies can be gauged from Table XXXII and the growth of assets and certain investments in Diagram IV, IVa & IVb.

TABLE XXXII

LIFE COMPANIES ESTABLISHED WITHIN GT. BRITAIN
PERCENTAGE DISTRIBUTION OF ASSETS *

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Year	Total Assets £millions	Mortgages	Loans on Public Rates, Municipal and County Securities, and Public Boards in U.K.	Loans on Policies and Personal Security	British Government and British Government Guaranteed Securities	Other Commonwealth Government Provincial and Municipal Securities	Foreign Government Provincial and Municipal Securities	Debentures	Stocks & Shares: Preference and Guaranteed	Stocks & Shares: Ordinary	Total of Columns 7, 8 & 9	Land, House Property and Ground Rent, Office Furniture, Etc.	Life Interest and Reversions	Agents Balances, Outstanding Premiums, Outstanding Interest, and Accrued Interest	Cash and Stamps	Establishment Expenses, Purchase of Business, Etc.
		%	%	%	%	%	%	%	%	%	%	%	%	%	%	%
1880	160	44·6	12·3	5·5	3·3	5·2	2·7	7·1	5·8		12·9	5·7	1·8	3·0	2·9	·1
1913	551	20·7	5·6	5·9	1·1	8·1	8·1	24·9	5·5	3·5	33·9	8·5	2·1	3·8	2·0	·3
1921	804	11·4	4·5	3·7	35·0	7·8	6·5	9·9	2·8	3·1	15·8	6·2	1·2	5·4	2·3	·4
1936	1,655	11·4	7·1	3·1	21·6	6·6	4·9	16·2	8·3	9·8	34·3	5·5	·4	2·9	2·1	·1
1948	2,724	6·8	3·7	1·0	39·7	3·8	4·2	8·9	7·9	10·7	27·5	5·3	·2	4·6	3·0	·0
1954	4,093	10·3	2·3	1·2	28·7	4·0	4·2	12·7	7·4	14·4	34·5	7·3	·2	5·1	2·2	·1

* *Including assets other than the life fund assets.*

DIAGRAM IV

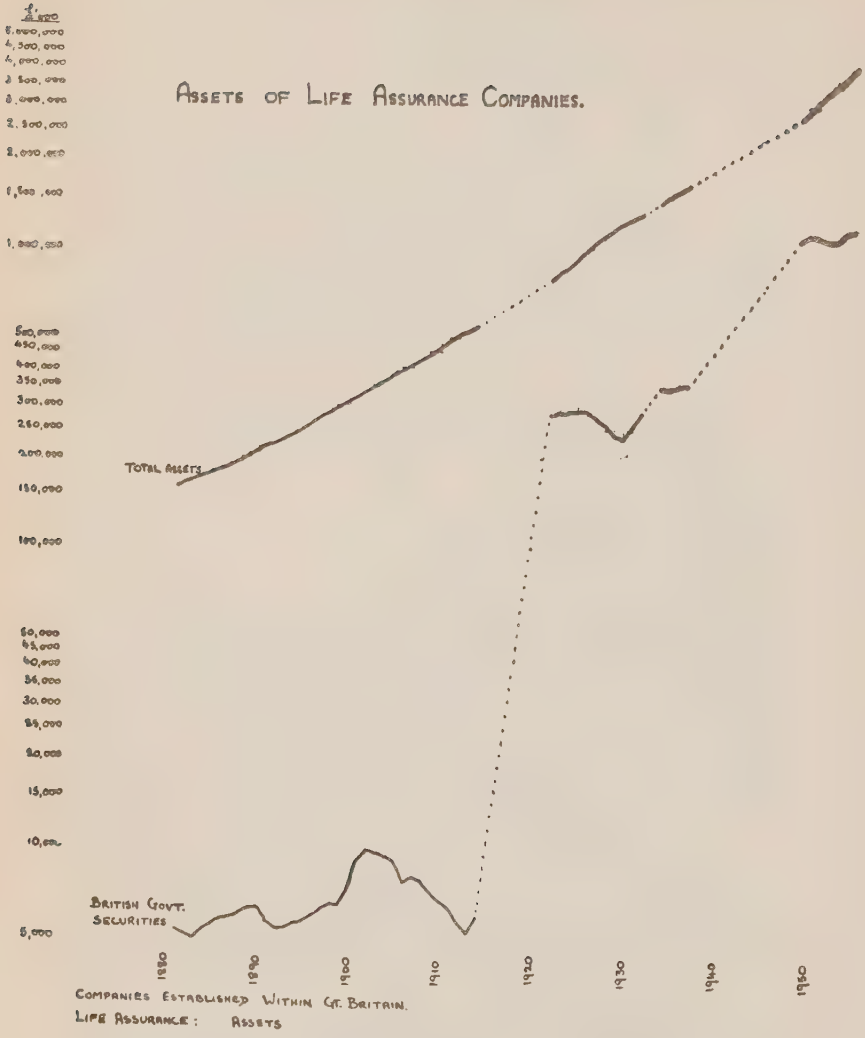

£'000
ASSETS OF LIFE ASSURANCE COMPANIES.
TOTAL ASSETS
BRITISH GOVT. SECURITIES
1880
1890
1900
1910
1920
1930
1940
1950
COMPANIES ESTABLISHED WITHIN GT. BRITAIN.
LIFE ASSURANCE: ASSETS

DIAGRAM IV a

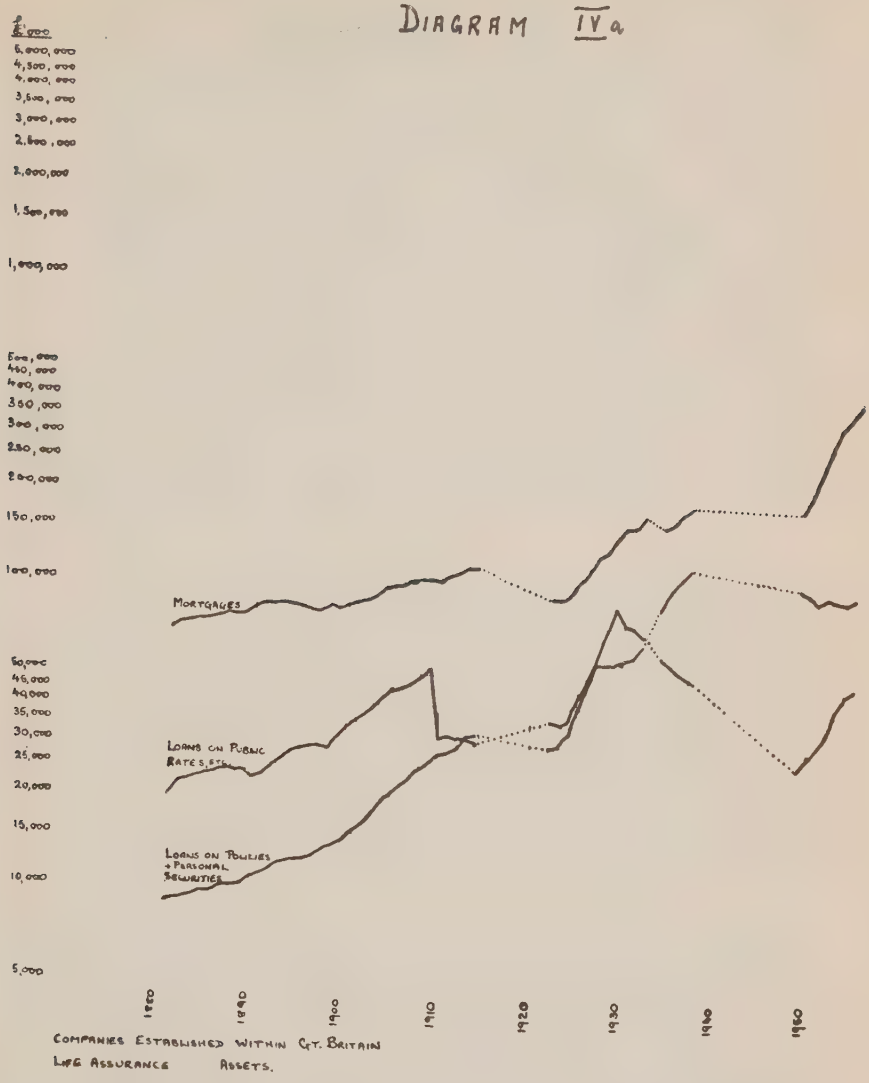

£'000
5,000,000
4,500,000
4,000,000
3,500,000
3,000,000
2,500,000
2,000,000
1,500,000
1,000,000
500,000
450,000
400,000
350,000
300,000
250,000
200,000
150,000
100,000
50,000
45,000
40,000
35,000
30,000
25,000
20,000
15,000
10,000
5,000
MORTGAGES
LOANS ON PUBLIC RATES, ETC.
LOANS ON POLICIES + PERSONAL SECURITIES
1850
1890
1900
1910
1920
1930
1940
1950
COMPANIES ESTABLISHED WITHIN GT. BRITAIN
LIFE ASSURANCE ASSETS.

DIAGRAM IV b.

COMPANIES ESTABLISHED WITHIN GT BRITAIN
LIFE ASSURANCE: ASSETS.

In 1880 investment in mortgages was the most popular item in the aggregate portfolio of British Life companies. A persistent decline followed until 1928-31. After this short interval of upturn there was a further decline, until in 1948 only 6·8 per cent of the total assets was invested in mortgages compared with 44·6 in 1880. The last year, 1954, shows a recovery to 10·3 per cent. A similar pattern of mortgage investment was followed in the United States. It is very difficult to give adequate reasons for the decline in the popularity of mortgage investment. If one could examine the portfolios of each company throughout the years more exact reasons could possibly be found. In 1954 one well known British Life company had 43·8 per cent of its total assets invested in mortgages. Probably the fact is that the development of cities was creating a strong demand for loans in the 1880's but that the total funds of the companies grew at a steady rate over the intervening years to 1954 and the demand for city loans diminished relatively.

It is interesting to trace the varying relative importance of British Government securities in the aggregate investment portfolio. In 1880 only 3·3 per cent of assets was invested and no marked change is seen until 1921 when the percentage rises to 35 per cent. A further decline occurs between the wars to a figure of 21·6 per cent in 1936. By 1948 this has risen to 39·7 per cent but another decline then ensues until 1954 is reached with the percentage at 28·7. Obviously the first and second world wars accounted for the very large percentage and absolute increases in this class of security.

Insurance companies' portfolios of British Government securities have been maintained at a higher level in Gt. Britain than in U.S.A. and Canada since the nationalisation of the major industries and the consequent issue of government-guaranteed securities. In the year 1945 over 45 per cent of U.S.A. Life companies' assets were invested in Federal Government securities. This percentage had declined to less than 10 per cent in 1954.

When one considers the changes in the holdings of British Government securities and the enormous increase from 1880 to 1954, it is thought that this is not the result of a deliberate policy on the part of the life office investment executive. It is much more likely to be the consequence of sheer force of circumstances. If the life offices had not chosen government securities, what else could they have done ? The extent of the problem faced by the investment executive is revealed by the figures in Table XXXIII.

TABLE XXXIII

INCREASE IN FUNDS

£million

1880—1890	...	...	58	cumulative
1890—1900	...	...	103	161
1900—1910	...	...	170	331
1910—1921	11 years	...	313	644
1921—1931	...	...	460	1,104
1931—1936	5 years	...	391	1,495
1936—1948	12 years	...	1,069	2,564
1948—1954	6 years	...	1,369	3,933

It is, in fact, said that when the broker of the largest life company moves on to the floor of the Stock Exchange the same expectancy awaits him as if he were the government broker !

Whilst government securities naturally assume an important place in life office investment portfolios the claims of commerce and industry have by no means been overlooked. In the 1880's 12·9 per cent of assets was invested in debentures, stocks and shares. The investment percentage increased steadily until the first world war, reaching 34·3 per cent in 1912.

After the war the percentage emerged at 15·8 per cent and rose steadily to 1929 reaching 31 per cent. With the economic depression, a minor decline brought it to 29·5 per cent. Again, the steady increase brought it to 34·3 per cent in 1936. The first figures available after the second world war showed a decline to 27·5 per cent but an all time high was reached in 1954 with a percentage of 34·5 per cent. Great interest is taken in these days in the activities of life offices in the equity market. The year 1909 gave the first separate figure for equities in the Board of Trade Summaries. In that year 3·9 per cent of total assets found investment in equities whilst British Government securities stood at 1·5 per cent ! A small reduction of 0·4 per cent brought us to 1913, another reduction of the same amount found equities at 3·1 per cent in 1921. From that point equities hardly ever looked back. There was a fall from 7·1 to 6·7 per cent in the depression, but the year 1948 showed 10·7 per cent rising to the highest figure recorded of 14·4 per cent in 1954. Clearly British life offices have found equities to their liking.

If we examine the portfolios of some of the most conservative life offices, we find an increasing interest in equities. Examining the breakdown of some leading offices for the years 1949 and 1954 we find an increase in percentage of total assets invested in equities in every single office. Some of the most outstanding offices having holdings of up to 29 per cent in equities alone. If we add the total holdings in debentures, stock and shares, we find holdings of up to 55 per cent of the total assets invested in commerce and industry.

TABLE XXXIV

LIFE COMPANIES ESTABLISHED WITHIN GT. BRITAIN
DISTRIBUTION OF ASSETS (13 Leading Companies) IN 1954

Company	Assets of Company £millions	1	2	3	4	5	6	7	8	9	10	11	12	13	14
		Mortgages	Loans on Public Rates, Municipal and County Securities, and Public Boards in U.K.	Loans on Policies and Personal Security	British Government and British Government Guaranteed Securities	Other Commonwealth Government Provincial and Municipal Securities	Foreign Government Provincial and Municipal Securities	Debentures	Stocks & Shares: Preferance and Guaranteed	Stocks & Shares: Ordinary	Total of Columns 7, 8 & 9	Land, House Property and Ground Rent, Office Furniture, tc.	Life Interest and Reversions	Agents Balances, Outstanding Premiums, Outstanding Interest, and Accrued Interest	Cash and Stamps
Co-operative	148	19·6	4·0	·6	38·8	·5	·5	3·5	7·2	14·2	24·9	8·4	—	1·6	1·1
Equitable Life	27	11·7	4·2	·9	19·2	3·2	·5	20·1	6·8	16·9	43·8	8·1	·5	7·7	·7
Legal & General	230	18·1	·9	·6	17·7	4·3	·2	9·5	4·2	8·8	22·5	28·9	2·5	2·5	1·8
London Life	38	11·4	2·5	3·2	36·0	3·2	·4	9·1	5·9	16·4	31·4	9·8	·0	1·9	·0
Prudential	772	3·4	2·0	1·0	39·8	4·2	·4	14·8	7·0	14·3	36·1	9·9	·0	2·4	·6
Scottish Equitable	19	6·5	·9	2·2	22·9	·2	·4	13·2	10·1	24·1	47·4	1·0	—	17·5	1·0
Norwich Union	119	43·8	·5	3·2	13·5	5·1	1·4	5·3	1·6	10·3	17·2	12·1	—	2·1	1·1
Sun Life ...	105	23·8	3·8	2·4	28·2	1·2	·0	26·4	3·2	7·2	36·8	1·7	·0	1·7	·5
Standard ...	144	6·4	·2	1·0	22·6	9·0	·1	22·5	10·6	19·9	53·0	·7	·0	6·4	·6
Pearl ...	220	1·5	2·1	1·0	39·9	6·5	1·4	14·3	7·3	17·7	39·3	5·3	—	2·1	·9
Scottish Widows	75	2·0	3·1	1·6	32·3	·1	1·5	10·3	15·3	29·5	55·1	2·8	·1	1·5	·0
Commercial Union	152	4·6	1·7	·7	19·4	4·5	12·7	15·1	8·5	11·4	35·0	5·4	·0	9·5	6·4
Refuge ...	139	13·3	5·0	1·4	36·5	3·4	·1	13·2	12·9	10·6	36·7	1·3	—	1·9	·4
All Companies	4,093	10·3	2·3	1·2	28·7	4·0	4·2	12·7	7·4	14·4	34·5	7·3	·2	5·1	2·2

This table analyses the total investments of the Companies concerned. Thus some companies may be doing a large amount of non-life business. In the case of the Commercial Union, had only assets applicable to life business been included, the combined percentage relating to Agents Balances, etc. and Cash would have been about 2·3% and foreign securities 0·5%.

To some persons this invokes a feeling of alarm for the stability of the companies. Others feel that insurance companies may attempt to control the policy of large industrial enterprises. When it is considered that this kind of investment is not a new policy of the assurance companies, and that most companies have enough to do to keep abreast of a very competitive market, it is hardly likely that any efforts will be made by assurance companies to control British industry, or for that matter to put too many eggs in one basket.

It is interesting to contrast policy in equity investment by United Kingdom life offices in the policy of U.S.A. life offices. Unlike American offices, British offices have not suffered from any degree of national or local government control in their investment policy. Investments of U.S.A. life offices are controlled by the State in which the company is domiciled. New York State is the most importance centre of life business. Up to 1905 considerable freedom of investment policy was permitted. This resulted in some considerable abuse. An investigation was sponsored by the New York State[1] which resulted in investment in equities becoming illegal. In 1928 New York life companies were permitted to invest 2 per cent of their assets in preferred stocks and by 1950 the figure had risen to 10 per cent for preferred stocks but nothing being still permitted for equities.[2]

Accounting practices at present in force in the United States are the deterrent, apart from individual State legislation, against investments in stocks and shares! Equity and debenture investment must be carried at market values. Fluctuations in such holdings would be reflected in surplus funds and if a sustained depression occurred the result may well be dangerous from a public confidence aspect. A company with more than 5 per cent of its assets invested in equities is exceptional, and for life business as a whole 4 per cent is invested in stocks and shares.

[1] Armstrong Investigation, 1906.

[2] A. J. McAndless and F. J. McDiarmid : Investments of Life Assurance Companies in U.S.A. from Proceedings of the Centenary Assembly of the Institute of Actuaries, 1950.

Table XXXV reveals the pattern of U.S.A. life assurance companies investments. The strong preference for mortgages and industrial bonds would appear to be necessary when so many States have legislation against investment in equities.

VII. *Life Assurance and Taxation*

The growth of life assurance has been described above (page 111). Although it is not possible to make a quantitative assessment, it seems fairly clear that a major factor influencing this growth has been the taxation relief afforded to policy holders and companies in the United Kingdom.

The main stimulant to business must have come from income tax relief. Since 1799 except between 1842-53[1], income tax relief has been granted to policy holders. These reliefs are well known but have varied from time to time. Whatever the variation they have always been an attraction to the prospective policy holder. At present the rate of relief will depend on the date of the policy and whether any part of the premiums paid falls within various legal restrictions. The insurance contract must be on the life of the taxpayer or his wife. The premiums must be paid by the taxpayer or his wife. The relief allowed in 1956-57 is two-fifths of allowable premiums as allowance from taxable income. If, however, the total premiums are under £25, the amount deductible from income is the amount of the premiums or £10, whichever is the lower sum.

The contributions paid to some types of pension scheme arranged with an insurance company also rank for tax relief on the same basis, and on other types of insured pension scheme higher tax reliefs are granted on the premiums.

In 1882, a statute was enacted which is known as the Married Women's Property Act. In 1894, Sir William Harcourt introduced his famous Estate Duty Act. Few people realised at the time that in these two Acts was another stimulant to the growth of British life assurance.

If an individual effects a contract of assurance on his own life and retains the benefit of the contract for himself or for his

[1] No income tax was in force for these years.

TABLE XXXV

DISTRIBUTION OF ASSETS
U.S. LIFE ASSURANCE COMPANIES
PERCENTAGE

	Total Amount 000,000 omitted	U.S. Government Securities	State and Local Bonds	Foreign Government Bonds	Rail Road Bonds	Public Utility Bonds	Industrial Bonds	Stocks	Mortgages	Real Estate	Policy Loans	Other Assets	Total
		%	%	%	%	%	%	%	%	%	%	%	%
1917	5,941	1·2	5·6	2·8	30·5	1·9	·8	1·4	34	3·0	13·6	5·2	100
1920	7,320	11·3	4·8	2·3	24·3	1·7	·7	1·0	33·4	2·3	11·7	6·5	100
1930	18,880	1·7	!5·4	·9	15·5	8·6	1·9	2·8	40·2	2·9	14·9	5x2	100
1940	30,802	18·7	!7·8	1·0	9·2	13·9	5·0	2·0	19·4	6·7	10·0	6·3	100
1945	44,797	45·9	!2·3	2·1	6·6	11·6	4·3	2·2	14·8	1·9	4·4	3·9	100
1955	90,432	9·5	!3·0	·4	4·3	15·5	20·1	4·0	32·6	2·9	3·6	4·1	100

Source : The Spectator and Institute of Life Insurance, U.S.A.

estate on his death, then the proceeds of the policy are liable to estate duty at the same *ad valorem* rate as the rest of the estate suffers.

If, however, an individual insures his life but does not retain the interest in the policy or policies, then subject to certain conditions either no duty is payable on the proceeds, or a very much reduced rate of duty than would have otherwise applied.

We do not know how much business has been transacted under the provisions of the Married Women's Property Act, 1882.

Section 33 of the Finance Act, 1954 made some curtailment to the scope of death duty exemption under Section 4 of the Finance Act of 1894, but nevertheless considerable concessions are still applicable. The concessions can be considerable and must provide an immense attraction to the far seeing and provident prospective investor in life assurance.

Supposing a person died on December 31st, 1954 and included in his estate were the following policies in which he never had an interest within the meaning of the Act of 1894.

Policy 1 for £2000 expressed to be for the benefit of his wife
„ 2 „ £1500 „ „ „ „ „ „
„ 3 „ £1500 „ „ „ „ „ „

£5000

Fixing rate of duty at 2 per cent.

Similar policies are also written for each of his four children. If it were not for the scope offered by the Married Women's Property Act, 1882, the whole of the £25,000 would be added to his general estate which we will assume attracted an estate duty tax of 40 per cent. This means that the policies would have suffered tax amounting to £10,000. Whereas in effect the duty leviable amounts to only £500. In addition to this income tax relief would have been granted on the premiums paid and prior to the Finance Act, 1954 no duty whatsoever would have been paid.

It may be said in passing that death duties on policies are not payable in such cases as a wife or husband insuring respective lives, providing the premiums are paid from each individual's own private financial resources. If we, therefore, add together the attractions of income tax relief and death duty exemption, it is possible to imagine the sales potential in the hands of the skilled professional vendor of life assurance.

For some years it has been obvious that the growth of the business of certain life offices has been tremendously influenced by the taxation relief gained by a company or institution which decides to institute a pension scheme. Pension schemes need not be serviced by a life office but in many cases they are arranged through an insurance company. When the additional benefit of life assurance is attached to pension benefits, the reward to any staff is very attractive. That all governments have continued to allow the whole of the employers' contributions to pension schemes to be regarded as a deduction in computing company taxation leads one to conclude that the growth of life companies specialising in this field will be accelerated in the near future. Even greater acceleration may be present when the effect of the 1956 Finance Act provisions with regard to annuities is known.

In many instances pension schemes which are linked with life assurance policies are free from death duties. Much of course depends on the provisions of the scheme and the manner in which the policies are written. Whether the motivating factor is taxation exemptions or a tremendous increase in a feeling of responsibility for employees welfare, the modern insurance undertaking has benefited greatly from the large increase in pension schemes business as is evidenced by the figures in Table XIII on page 130 above.

Having provided the policy holder with substantial attractions to invest in life assurance, successive British governments have lent a sympathetic ear to the pleas of life offices where dangers could be foreseen from the companies taxation viewpoint.

The subject of income tax, surtax and profits tax whether with regard to an individual person, firm or company is complicated. Where specialising financial institutions like assurance companies are concerned it is even more complex.

Because another stimulant to growth is concerned we outline below the position with regard to the taxation of a life office. Precise measurement of the stimulus provided is impossible.

Ordinary life assurance business is conducted by three types of organisation. The first type is called a " composite " office and in it are executed all manner of insurance contracts such as fire, accident, marine and life. The second kind of office conducts only life and annuity business and it is owned, like the first, by the shareholders. The third type of organisation conducts only life and annuity business but it is owned by its policy holders and is termed a mutual office.

The Income Tax Act 1952, Section 426, lays down the rule that life assurance business profits must be separately assessed in all kinds of assurance companies. Prior to 1915 a composite office was assessed on its profits from all sources of business. Section 427 (1), Income Tax Act, 1952, states that any proportion of the profit of a life company reserved for policy holders shall be exempted from income tax. Most life companies allocate at least 90 per cent of profits to with-profit policy holders. It is possible to calculate that substantial gains in the incidence of taxation are available for the tax payer who can invest in with-profit policies. Distribution of profits to shareholders is taxed both in income tax and profits tax legislation. Insurance companies are taxed in effect on the income derived from investments less expenses and commissions incurred in conducting the business of the company. This basis is subject to one further important relief : introduced as an emergency concession by the Finance Act of 1940. The concession amounts to relief on the difference between the standard rate of income tax and 7s. 6d. calculated on gross investment income less commission, expenses and sums allocated to shareholders.

Thus, the value of this concession to a company with £100 million invested in the Ordinary Branch earning £5 million gross investment income might be £175,000.

Gross Investment Income ...		5,000,000
Less Commission and Expenses	1,000,000	
Shareholders	500,000	
		1,500,000
		3,500,000

which at 1/- in £ amounts to £175,000.

It is thought that where the rate of tax relief is significant, say between one shilling and two shilling and sixpence in the £, the fact must be important when bonuses are declared in subsequent years.

Proprietary companies are also liable to profits tax in respect of profits allocated to policy holders. It is thought that this point is of little significance when compared with the general stimuli offered through taxation.

The Finance Act, 1956, Section 24 provides a new basis for taxing life assurance companies carrying on pension annuity business. The section is in effect carrying out the recommendations of the Millard Tucker Committee on the Taxation Treatment of Provisions for Retirement. Under this section the annuity funds of assurance companies are relieved from income tax so far as they are the invested premiums arising from deferred annuity business for the self-employed.[1] It is, of course, early days to venture an opinion on the future of this kind of business and no part of the present growth of life assurance companies can be yet attributed to this concession.

In conclusion it seems fairly clear, therefore, that the taxation relief afforded to policy holders and companies has been a major factor in stimulating the growth of life assurance in the United Kingdom. And with the current high levels of taxation any further reliefs might be an even stronger stimulant to growth.

[1] Income tax relief extends to superannuation funds approved under Section 379 of Income Tax Act 1952 (see Section 24(7) (b)).

VIII. *Summary*

The number of life assurance companies has remained fairly stable throughout this century, but the volume of business done has increased steadily. Premium income in the ordinary branch increased by 4·4 per cent per annum between the wars and in the industrial branch by 3·7 per cent. Since 1946 premium income has been increasing at over 10 per cent per annum in the ordinary branch and at over 4 per cent in the industrial branch. Total premium income has reached a level of just over 3 per cent of total personal income in U.K. which is very close to the corresponding figure for U.S.A. of 3·1 per cent. Sum assured per unit of premium income, however, is almost twice as high in U.S.A. as in U.K. This seems to be a reflection of two factors, namely, the emphasis in U.S.A. on the selling of insurance (i.e., cover) *per se* compared with the emphasis in U.K. of combining *saving* with *insurance* via endowment policies, and the greater development of " schemes " business in U.S.A. (group life, credit, life, family income policies, etc.).

Within the British overall growth, the three outstanding features are :—

(i) The steady drift from with-profits to without-profits policies.

(ii) The increasing importance of endowment policies, and

(iii) the rapid development of " schemes " business since the last war.

In the ordinary branch the share of the total business handled by the major companies has increased over the century, but while the Prudential has remained the largest company throughout there has been a substantial amount of jockeying for position in the remaining top places. It seems that the rapidly growing companies have grown by expansion of business rather than by absorption or association with other companies. In recent years companies specialising in "schemes" business have shot rapidly ahead.

The examination of the variation in expense ratio with size of company is complicated by factors such as the type of business done and the relative amounts of new business, but it would appear that there is a tendency for the larger companies to have lower expense ratios.

There are substantial variations between companies in mortality experience, interest rates earned and as a consequence in premium rates charged. Substantial variations also occur in the rate of profit, but there is some evidence of a decline in profit rate with size of company.

Total assets of British life companies available for investment have grown from £160 million in 1880 to £4,093 millions in 1954. Life assurance companies are among the leading investors in Stock Exchange securities. In 1954 nearly 29 per cent of their assets were invested in British Government and British Government Guaranteed securities. Commerce and industry, however, claimed the largest share of investment obtaining 34·5 per cent of total assets. In the United States the proportion of life assurance companies assets invested in commerce and industry is 28 per cent in 1954. No evidence has been found of financial instability even though there is no government regulation of investment policy of British life offices. This contrasts strongly with the position in the United States where state regulation has been in force since 1906 following unsatisfactory financial administration.

Taxation allowances to policy holders probably have been the major stimulant in the growth of life companies. Since 1940 life companies have enjoyed a taxation relief which depends on the proportion of profits allocated to policy holders and the amount by which the standard rate of income tax in force exceeds 7s. 6d. in the £. The recent development of pension schemes through life companies is also greatly stimulated by taxation concessions both to employer and employed. Finally the Finance Act, 1956 relieves annuity funds of insurance companies from income tax in so far as they are the invested premiums arising from the deferred annuity business for the self-employed.[1]

[1] See footnote, page 169.

APPENDIX 1

Accounts of Assurance Companies and the determination of Profits or Surplus.

United Kingdom assurance companies have long enjoyed a comparative freedom from State control and intervention. Statutes have been in force from 1870 which governed certain aspects of their activities. In 1909 the principal act was passed and is known as the Assurance Companies Act of that year. It repealed the Life Assurance Companies Acts of 1870, 1871 and 1872. In 1946 a further Act was passed which is known by the title Assurance Companies Act, 1946. It may be noted in passing that the 1948 Companies Act affects the yearly financial accounts in a more modified manner than it does those of ordinary industrial and commercial concerns.

All acts from 1870 have required the deposit of accounts with the Board of Trade. The accounts must be drawn up in the manner prescribed by the relevant act. Without the data which have been deposited with the Board of Trade since 1870 it would have been impossible to prepare a paper of this type. The annual accounts of every British life assurance company have been filed since 1870.

The form of accounts of life assurance companies is prescribed in several Schedules. We are mainly concerned in our study with the material available in some of the accounts required under the First Schedule (A), Second Schedule, Third Schedule and Fourth Schedule. For reference the main forms required are reproduced below but without all the notes usually appended.

FIRST SCHEDULE

(A)—FORM APPLICABLE TO LIFE ASSURANCE BUSINESS

Revenue Account of the for the Year ending in respect of Life Assurance Business.

	Business within the United Kingdom.	Business out of the United Kingdom.	Total.		Business within the United Kingdom.	Business out of the United Kingdom.	Total.
	£ s. d.	£ s. d.	£ s. d.		£ s. d.	£ s. d.	£ s. d.
Amount of life assurance fund at the beginning of the year	—	—		Claims under policies By death ... ... By maturity ...			
Premiums :—				Surrenders, including surrenders of bonus			
Consideration for annuities granted ...				Annuities			
£ s. d.				Bonuses in cash ...			
Interest, dividends, and rents ...				Bonuses in reduction of premiums ...			
				Commission ...			
				Expenses of management	—	—	
Less income tax thereon	—	—		Other payments (accounts to be specified)	—	—	
				Amount of life assurance fund at the end of the year, as per Third Schedule ...	—	—	
Other receipts (accounts to be specified)	—	—					
£				£			

Section 4

SECOND SCHEDULE

PROFIT AND LOSS ACCOUNT OF THE FOR THE YEAR ENDING 19

	£ s. d.		£ s. d.
Balance of last year's account		Dividends and bonuses to shareholders	
Interest and dividends not carried to other accounts	£ s. d.	Expenses not charged to other accounts	
Less income tax thereon ...		Loss realised (accounts to be specfied)	
Profit realised (accounts to be specified)		Other payments (accounts to be specified) ...	
Other receipts (accounts to be specified) ...		Balance as per Third Schedule	
£		£	

Section 4.

THIRD SCHEDULE

Balance Sheet of the..... on the 19

LIABILITIES	£ s. d.	ASSETS	£ s. d.
Shareholders' capital paid up (if any) ... ... £		Mortgages on property within the United Kingdom	
Life assurance funds :—		Do. do. out of the United Kingdom	
Ordinary branch		Loans on parochial and other rates	
Industrial do.		Do. Life interests	
Annuity fund		Do. Reversions	
Sinking fund and capital redemption fund		Do. Stocks and shares	
Profit and loss account		Do. Company's policies within their surrender values ...	
Other funds (if any) to be specified		Do. Personal security	
		Investments :	
		Deposit with the High Court (securities to be specified) ...	
Claims admitted or intimated but not paid		British Government securities	
Life assurance		British Government Guaranteed Stocks	
		Municipal and county securities, United Kingdom	
Annuities due and unpaid		Indian and Colonial Government securities	
Other sums owing by the Company (to be stated separately under each class of business).		Do. provincial securities	
		Do. municipal do.	
		Foreign Government securities	
		Do. provincial securities	
		Do. municipal do.	
		Rent charges	
		Freehold ground rents	
		Leasehold do.	
		House property	
		Life interests	
		Reversions	
		Agents' balances	
		Outstanding premiums	
		Do. interest, dividends, and rents	
		Interest accrued but not payable	
		Bills receivable	
		Cash :	
		On deposit	
		In hand and on current account	
		Other assets (to be specified)	
£			£

(FORM REFERRED TO UNDER HEADING No. 7 IN FOURTH SCHEDULE (A).)

Valuation Balance Sheet of as at 19.

Dr.	£	Cr.	£
To net liability under Life Assurance and Annuity transactions (as per summary statement provided in Fourth Schedule (A))		By Life Assurance and Annuity funds (as per balance sheet under Schedule 3)	
To surplus, if any...		By deficiency, if any	

From an overall consideration it will be realised that when transactions are recorded in the above manner the effect is that the First Schedule shows an analysis of revenue and expenditure, together with the opening and closing balances. The Second Schedule shows a further analysis on the same lines. The Third Schedule reveals how the residual balance in the First and Second Schedules has been invested. The Fourth Schedule throws out a positive or negative difference when the fund on the First Schedule is weighed against an actuarial valuation of the possible claims on the life fund. It has been stated previously that very little State supervision has ever been exercised over insurance funds in this country. The result is that a conservative doctrine of accounting predominates in the insurance field. Providing the auditor and actuary certify that the assets exceed the possible claims upon them all is well. We are told by experts that the actuarial valuation of the surplus is made to suit the bonus policy of any particular company. From an ordinary commercial or economic viewpoint we can assume the surplus figures to be well understated.

In addition to the understatement of the surplus, transfers are made from the First Schedule to reserve and provision accounts. The aggregate or sum of the yearly reserve transfers is rarely revealed on the Third Schedule or Balance Sheet. It is deducted from the assets held to represent the Life Fund. It is therefore virtually impossible for the outsider to judge the efficiency of the company from a profits and capital employed aspect. On the question of assets and liabilities, it is clear that generally the assets are understated and the liabilities overstated. We took the trouble to total the transfers to reserve which had been made by a well known company over a period of seventy years. They amounted to £25 million. Only a few millions were shown in revealed reserve funds. The assets of the company appeared on this count alone to be underestimated by 12 per cent.

Several leading companies have very kindly supplied us with details of their transactions over the years 1909-1955. The following notes on surplus of one company are reproduced as an illustration of the arbitrary element in the determination of surplus.

YEAR 1921.

1. Valuation made on less stringent basis than those at the end of the two previous quinquennia.

YEAR 1931.

2. Valuation made on less stringent basis than those at the end of the two previous quinquennia. Had the previous basis been used the surplus would have been reduced by £464,860.

YEAR 1936.

3. Valuation made on more stringent basis than that used at end of previous quinquennium. Had the previous basis been used the surplus would have been increased by £1,039,690.

YEAR 1941.

4. Valuation made on more stringent basis than that used at end of previous quinquennium. Had the previous basis been used the surplus would have been increased by £1,337,639.

YEAR 1946.

5. Valuation made on more stringent basis than that used at end of previous quinquennium. Had the previous basis been used the surplus would have been increased by £3,071,230.

It therefore must be admitted that from the aspect of measuring growth in profit the published surplus figures of United Kingdom life assurance companies are not very enlightening.

As an illustration of the actual figures we show in Diagram V the life fund and the profits of three leading companies. It is thought that the graphs might reveal some connection between each company and their life fund and surplus policy. The graphs do show that the common catastrophe of war has a marked effect on life fund valuation for surplus policy. Each company shows a common upswing in surplus a few years after a period of war.

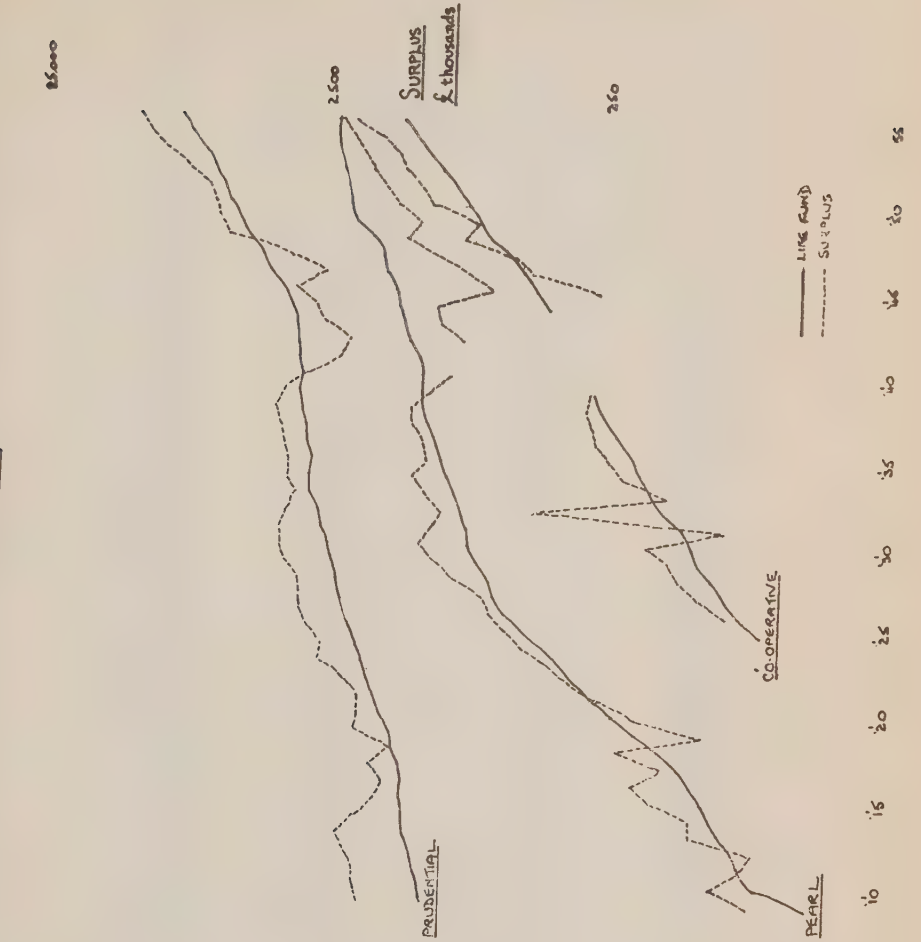


DIAGRAM V
LIFE FUND £ millions
SURPLUS £ thousands
PRUDENTIAL
CO-OPERATIVE
PEARL
LIFE FUND
SURPLUS

APPENDIX II

TABLE I

LIFE ASSURANCE—ORDINARY BRANCH

(Companies established within Great Britain)

Year	Premiums £mn.	Sums Assured £mn.	Life Fund £mn.
1881	12·5		127·7
1882	12·8		131·9
1883	12·9		135·6
1884	13·2		139·5
1885	13·4		142·8
1886	13·7	422	146·1
1887	14·4	439	149·9
1888	15·0	443	154·9
1889	15·5	444	160·5
1890	16·0	456	165·9
1891	15·7	481	170·5
1892	17·9	492	176·4
1893	18·3	501	183·0
1894	19·4	530	190·2
1895	21·0	553	199·3
1896	21·9	564	209·2
1897	22·2	590	219·8
1898	23·2	603	228·9
1899	23·4	619	238·2
1900	23·5	641	246·1
1901	24·4	663	255·3
1902	25·0	679	265·0
1903	25·8	688	274·2
1904	27·3	700	284·5
1905	27·5	715	296·0
1906	28·1	730	306·4
1907	28·9	760	315·4
1908	30·0	771	325·9
1909	31·9	785	336·4
1910	31·3	803	348·6
1911	31·9	830	359·8
1912	32·8	839	370·5
1913	33·9	858	381·5
1914	34·3	873	390·1
1915	33·8		390·5
1916	34·0		395·2
1917	34·5		396·4
1918	38·6		411·0
1919	43·5		425·9

Table I. Life Assurance—Ordinary Branch. (Companies established within Great Britain)—*continued.*

Year	Premiums £mn.	Sums Assured £mn.	Life Fund £mn.
1920	47·5		437·5
1921	49·0		456·6
1922	51·5	1,149	479·5
1923	54·3	1,187	506·3
1924	58·8	1,214	532·7
1925	65·8	1,296	570·1
1926	65·4	1,359	605·6
1927	71·4	1,404	640·4
1928	76·0	1,445	674·3
1929	78·0	1,498	710·1
1930	75·5	1,598	734·5
1931	75·4	1,676	754·3
1932	86·4	1,715	793
1933	90·1	1,769	833
1934	91·9	1,821	868
1935	95·6	1,950	903
1936	98·1	2,179	943
1937	99·7	2,161	984
1938	102·1		1,023
1939	99·5		1,042
1940	96·4		1,054
1941	95·5		1,081
1942	99·8		1,115
1943	103·6		1,153
1944	109·5		1,193
1945	117·5	2,631	1,244
1946	134·5	2,855	1,313
1947	158·9	3,045	1,391
1948	173·0	3,244	1,478
1949	185·8	3,521	1,573
1950	203·9	3,930	1,675
1951	232·4	4,487	1,786
1952	253·6	4,821	1,914
1953	274·7	5,143	2,078

Source : Summaries of Statements of Assurance Business deposited with the Board of Trade.

(Premiums include considerations for Annuities and Sums Assured the amounts of Annuities).

TABLE II

INDUSTRIAL ASSURANCE £'000

Year	Companies			Societies			Total		
	Sums Assured	Pre-miums	Fund	Sums Assured	Pre-miums	Fund	Sums Assured	Pre-miums	Fund
1914	428,691	17,983	58,701						
1915		18,565	61,382						
1916		19,573	64,992						
1917		20,959	69,435						
1918		22,398	74,028						
1919		25,350	80,519						
1920		29,269	89,711						
1921		31,088	100,280						
1922	693,454	31,580	111,451						
1923	715,603	33,213	123,200						
1924	750,716	34,133	136,420		8,130	29,322		42,263	165,742
1925	832,433	36,615	151,785	199,312	8,567	31,846	1,031,745	45,182	183,631
1926	870,218	36,802	165,849		8,633	34,733		45,435	200,889
1927	903,161	38,805	179,143		9,246	37,651		48,051	217,107
1928	920,125	40,877	192,029		9,642	40,902		50,519	223,375
1929	952,820	42,002	203,113		10,026	43,606		52,028	247,042
1930	1,006,382	43,811	217,324	245,363	10,408	46,825	1,25,745 0	54,219	264,493
1931	1,035,205	45,407	226,258		10,670	50,434		56,077	276,693
1932	1,050,590	46,355	243,031		10,978	53,362		57,332	296,393
1933	1,088,770	47,663	259,262		11,275	56,925		58,938	316,187
1934	1,123,440	50,072	277,518		11,819	60,624		61,891	338,142
1935	1,176,405	51,503	296,712	279,286	12,223	63,903	1,455,691	63,726	360,615
1936	1,215,813	53,629	317,085		12,761	67,972		66,390	385,056
1937		55,792	337,250		13,442	72,401		69,235	409,651
1938		57,997	359,219		13,829	76,192		71,826	435,411
1939		59,835	374,370		14,335	80,697		74,170	455,067
1940		62,189	388,568		14,562	84,873		76,751	473,441
1941		63,557	406,981		15,034	89,788		78,592	496,768
1942		67,963	430,357		15,873	95,548		83,836	525,906
1943		72,186	455,974		17,007	102,374		89,193	558,349
1944		75,701	484,936		17,815	109,335		93,516	594,271
1945		79,392	526,851		18,408	116,553		97,800	643,404
1946		83,562	568,645		19,599	123,832		103,161	692,477
1947		91,161	604,130		21,215	132,517		112,376	736,640
1948		97,103	640,029		22,757	141,677		119,860	781,706
1949		100,334	671,114		23,777	150,362		124,111	821,476
1950		103,474	703,000		24,133	159,266		127,606	862,265
1951		108,211	724,854		24,874	167,173		133,084	892,027
1952		111,972	746,540		25,633	172,452		137,605	918,992
1953	2,337,910	116,641	786,488	543,788	26,424	182,647	2,881,698	143,066	969,136

Source : Annual Reports of the Industrial Assurance Commissioner and Statments of Assurance deposited with the Board of Trade.

J. Johnston and G. W. Murphy

Department of Economics,
Manchester University.

An Engel-Type Curve for Cash

The curve known by the name of Engel indicates the statistical association between the consumption of an individual commodity and the income of the consumer. The analogous curve for assets would indicate the association between the amount held of an individual asset and the wealth of the holder. This note describes a curve of the latter type for cash.

In economic parlance a commodity is said to be a luxury if the income elasticity of demand for it exceeds one and a necessity if it falls below that figure. It is not uncommon for a commodity to trace an Engel curve showing that commodity as having the property of a luxury at low incomes, that of a necessity at higher ones, and of reaching eventually a point of saturation (income elasticity zero).[1] By analogy, we might expect to find that Engel-type curves for (at least some) assets have similar characteristics, it being understood, of course, that wealth elasticity is now the appropriate criterion, not income elasticity.

In the case of cash there are strong *a priori* considerations which would lead one to believe that it would trace an Engel-type curve conforming to this pattern. For very low wealth groups the advantages flowing from the liquidity of cash must be set principally against the utility derived from essential household equipment,[2] and, as compared with the latter, cash might well be a luxury. Once past the very low ranges of wealth, however, basic assets (such as house furniture for example) are in hand, and the principal alternatives against which the liquidity of cash has to be weighed are themselves luxuries : at one wealth range it may be a small second-hand car ; at another bonds or shares. On that account, even if a luxury for very low wealth groups, cash would be expected to drop back to a necessity as we go up the scale.

[1]See S. J. Prais, "Non-Linear Estimates of the Engel Curves," *Review of Economic Studies*, Vol. XX (1952-53), pp. 87-104 ; J. Aitchison and J. A. C. Brown, "A Synthesis of Engel Curve Theory," *ibid.*, Vol. XXII (1954-55), pp. 35-46. Logically there is one more stage beyond saturation—that of inferiority (income elasticity negative). See J. R. Hicks, *Value and Capital*, Oxford : Clarendon Press, 2nd ed. 1946, pp. 27-29 ; H. Wold and L. Jureen, *Demand Analysis*, New York : Wiley & Sons, 1954, pp. 272-75.

[2]In the present context all durable consumer goods are treated as assets.

Moreover it would be expected that it would become progressively a more pronounced necessity, in the sense of the Engel-type curve exhibiting a continuously diminishing wealth elasticity, and eventuating in something like an asymptotic approach to a point of saturation (wealth elasticity zero). For on the one hand new luxuries are virtually inexhaustible—a mink, a Picasso, a bathing pool—whilst, on the other, the probability of the simultaneous occurrence of a number of events, against which the liquidity of cash is useful insurance, rapidly approaches zero, as, with increasing affluence, remoter contingencies are brought in.[1] (This relates, of course, to the precautionary motive for holding cash. The speculative motive will be taken account of later). Moreover, working in the same direction is the fact that the costs attached to buying bonds or shares grow proportionately less the greater the transaction, and such assets become therefore increasingly attractive relatively to cash as we move up the wealth scale.

Data from which Engel-type curves for assets can be constructed have now been published for many years in the annual reports of the Commissioners of Inland Revenue. They are collected in the course of administration of estate duty and therefore relate only to estates liable to such duty, *i.e.* those exceeding £2,000.[2] For ranges of wealth below the £2,000 mark there are unfortunately no published figures from which their asset distribution can be gauged. The reports of the Oxford savings surveys —the only hopeful source for such figures at present—do not as yet tabulate holdings of individual assets by wealth groups, though they do by income groups. From the present point of view this is rather unfortunate, for it means that the hypothesis

[1]This argument points to a limit in the growth of individual precautionary balances only if there is also a limit to the rise in the potential cost of each contingency as we move up the wealth scale. Only, for example, if both the man with £500,000 and the man with, say, £50,000 send their wives to Harley Street when they are a bit off colour. This footnote has been inserted at the instigation of Professor H. G. Johnson who has been most helpful in other ways, too.

[2]Estates up to £3,000 are exempt from duty if death has occurred after 29th July, 1954. But since the relevant table in the latest report (at the time of writing) relates to estates on which duty was paid in the year ending 31st March, 1955, and since there is usually a considerable interval between an estate becoming liable to duty and the actual payment of it, the alteration of the exemption limit affects but little the available figures.

that cash might be a luxury asset for very low wealth groups cannot be checked. But even without this, the data available are sufficiently interesting to merit a close look.

In Table I the columns of cash percentages have been taken from the last three Inland Revenue reports.[1] There is one such column for each of the financial years covered by these reports. From the cash percentages for any two adjoining wealth groups can be computed a crude arc wealth elasticity for

TABLE I

Range of net capital value of estate		Financial Year 1952-53		Financial Year 1953-54		Financial Year 1954-55	
		Average Cash per-centage	Arc Elasticity	Average Cash per-centage	Arc Elasticity	Average Cash per-centage	Arc Elasticity
Exceeding	Not Exceeding						
£	£						
2,000	3,000	20·42		21·83		24·55	0·92
3,000	5,000		0·67		0·79	23·62	0·82
5,000	10,000	16·02		18·72		21·22	
			0·88		0·72		0·64
10,000	15,000	15·08		16·25		17·76	
			0·45		0·28		0·30
15,000	20,000	12·53		14·25		14·04	
			0		0		–0·28
20,000	25,000	9·75		11·01		10·17	

[1]96th Report, H.M.S.O., Cmd. 9030, table 104 ; 97th Report, H.M.S.O. Cmd. 9351, table 105 ; 98th Report, H.M.S.O. Cmd. 9667, table 105. Cash here includes currency holdings, current and deposit accounts with commercial banks and deposits in the Trustee and Post Office savings banks. The last category is to a certain extent regarded also as investment by the depositors, and this introduces some bias to the story revealed by the tables. As for the commercial bank deposit accounts, they yielded an all but insignificant return in the early part of the period covered by the tables (allowing for the delay referred to in the previous footnote).

cash holding over that particular region of the Engel-type curve.[1] The results of such computations are shown in columns adjoining those with the cash percentages. Individually, none of these elasticities is significant. But taken together they reveal an interesting and consistent story.

It is a story much as expected. With but one exception, the elasticities are seen to fall continuously as we go up the wealth scale. The saturation point in the desire for cash holding duly appears between the ranges of £15,000-£20,000 and £20,000-£25,000; indeed it appears with a spurious exactitude, the elasticities in that region being a round zero (when taken up to two decimals) for two of the three years. And, whilst it is clear from this evidence that beyond the £2,000 mark cash is not a luxury, a backward projection of the elasticities of the last column would suggest that they might in fact rise above unity for very low wealth groups.

The data of Table I conform, then, to *a priori* expectations. This is doubtless satisfactory. But the evidence of factual data is charged with greater implications if it confounds, than if it confirms, accepted hypotheses. From this point of view what follows is the more important part of this note.

It will be observed that Table I ends with estates in the range of £20,000-£25,000. But the Inland Revenue tables from which the cash percentages have been extracted do not end there. Nor is the part which has been omitted unimportant. For of all the cash in estates on which duty was paid in any one of the years under consideration, more than a quarter was to be found in those exceeding £25,000. Take a look, therefore, at Table II which picks the story up again at the £20,000-£25,000 range of estate and carries it as far as it would go, though without any elasticity calculations. The surprising thing which it reveals is that

[1] Let $V_i \equiv$ middle of the ith range of net capital value of estate; and $p_i \equiv$ cash percentage corresponding to the ith range. Then: arc wealth elasticity of cash holding between the ith and $(i+1)$th ranges $= \frac{2(p_{i+1}V_{i+1} - p_iV_i)}{(p_i + p_{i+1})(V_{i+1} - V_i)}$, where $\frac{p_{i+1}V_{i+1} - p_iV_i}{V_{i+1} - V_i}$ is the increment of cash divided by the increment of wealth, and $\frac{2}{p_i + p_{i+1}}$ is the reciprocal of the average of the cash percentages corresponding to the two ends of the arc.

beyond the £20,000 mark the cash percentage either fails to decline at all, or, if it declines, it does so gently, and, more often than not, temporarily. In terms of our crude wealth elasticities this means that they hover for much of the time around unity, and to the extent that they depart substantially from it, they do so in both directions. And the picture is not significantly altered if we omit from consideration the last two or three ranges, as including too few estates to constitute adequate samples of the corresponding wealth groups.

TABLE II

Range of net capital value of estate		Average Cash Percentage		
		Financial Year 1952-53	Financial Year 1953-54	Financial Year 1954-55
Exceeding £000	Not Exceeding £000			
20	25	9·75	11·01	10·17
25	30	10·76	9·70	9·59
30	40	9·98	9·64	9·11
40	50	9·39	8·71	9·27
50	60	7·98	9·95	8·57
60	80	9·00	9·92	9·83
80	100	10·61	8·47	6·73
100	200	8·89	8·06	7·60
200	300	6·55	10·68	7·65
300	500	7·65	10·05	10·07
500	1,000	6·06	9·58	11·77
1,000	2,000	7·66	12·59	17·19
2,000		6·02	34·91	9·32

How is this phenomenon to be accounted for ? Three possible explanations will be offered here. In the first place, to avoid massive selling of illiquid assets in possibly unfavourable markets in order to pay estate duty, some provision might be made to increase liquidity in anticipation of death. Such provision would require the volume of cash to increase proportionately more than the value of the estate, since the rate of estate duty is progressive. This would then counteract the tendency of the cash percentage to fall. In this connection, it may not be entirely a coincidence that, as shown in the tables, the counteracting force only operates

beyond the £20,000 mark, whilst up to estates of £20,000 the rate of estate duty is below the average cash percentage. (Should this hypothesis have any validity, it would mean, needless to say, that it would impart a bias to the sample we are dealing with here, inasmuch as the life expectation of those who actually die in any given period must be shorter than that of the wealth group to which they belong as a whole).

Secondly, there is the effect of a progressive income tax to take into account. Beyond a certain point in the wealth scale the principal substitutes for cash would be income-yielding assets. But *after tax* the yield of such assets falls as their quantity increases. The fact, then, that the utility derived from cash must be set against that derived from assets not with a constant but with a diminishing yield, would also tend to counteract the tendency of the cash percentage to fall. If, in addition, this counteracting were weaker the lower the marginal rate of income tax paid by the asset holder, we would have an explanation for the fact that there is not much evidence of a counteracting force up to estates of £20,000.

The third hypothesis is not related to the effects of taxation but to those of the speculative motive for holding cash, of which we have not taken account so far. Speculation is bound to be unprofitable if engaged in a small scale, because the expenses of buying and selling would then form a considerable proportion of turnover. There is thus a point on the wealth scale below which the speculative motive for holding cash does not operate. But when it does begin to operate, it will clearly boost the average cash percentage. (The individual, as distinct from the average, cash percentage at any particular time will also depend on whether one is a "bull" or a "bear"). Should this, then, be the principal explanation for the break in the fall of the average cash percentage in the region of £25,000, it would follow that this figure could be identified as the minimum capital required before engaging in speculation.

Taking the matter one step further it would be possible to explain by the speculative motive not only the break in the fall of the cash percentage at a given point, but also the failure of the fall to resume its pronounced course beyond that point. For this it would be necessary to suppose (*a*) that an increase in the scale of speculation requires a more than proportionate increase in the average volume of cash held and (*b*) that the scale of asset holders' speculation is positively correlated with the size of their total capital.

JOHN SPRAOS

University of Sheffield.

POSTSCRIPT.—After this note had gone to the printers, the 99th Inland Revenue Report (H.M.S.O., Cmnd. 54) became available. The data which it gives in table 130 for the financial year 1755-56 conform to the pattern of the previous three years indicated above. This confirms that the pattern (which can also be traced back further than was done here) is a reasonably stable one.

Private Company Finance in the Post-War Period[1]

This article embodies the results of a field inquiry into how private companies have been financed in the inflationary conditions of the post-war period. In order to keep the data obtained within manageable bounds, and at the same time avoid the type of bias that might result from an arbitrary selection of firms scattered throughout the country, the inquiry was confined to a definite industrial area in the Midlands. The data were obtained by personal interviews with the owners of the firms willing to assist the inquiry, which amounted to about half of the firms approached. Information supplied varied from general statements to the provision of complete records. Questions were limited to the smallest number possible, as the people interviewed were busy men. Assurance was given that the information would be treated as confidential.

The Composition of the Sample.

The firms investigated fell into three main groups :—

(*a*) *Textile and Clothing :* comprising narrow fabric weaving, specialised weaving, sock and stocking manufacture, lace, underclothing, etc., 23 firms.

(*b*) *Engineering and Construction :* comprising small subcontractors, manufacturers of industrial equipment, of components for incorporation into other products, large and small contractors, etc., 23 firms.

(*c*) *Miscellaneous :* comprising a great variety of manufacturers not included in (*a*) and (*b*). These were not separately classified as in many cases there were one or two firms only in a particular industry and disclosure of figures for an industry would have meant possible

[1]This information was obtained for the purpose of a PhD. thesis of the University of London, on a topic suggested by Professor F. W. Paish. Both Professor Paish and Professor Brian Tew have offered valuable comments on the original draft (University of London Library) as well as assistance when the work was in progress. They are in no way responsible for the present article, which seeks to extract some of the more interesting material discovered.

disclosure of the firm. Garages and merchanting businesses are also included under this heading. Retail firms were not approached, 26 firms.

THE SIZE AND AGE OF THE FIRM.

Firms were divided into three sizes: small, medium and large. The division was by "inspection" as no other complete information was available as to capital, numbers employed, area occupied, turnover, etc. By small is meant the manufacturing firm with up to about twenty employees, occupying a limited space and with a small turnover. By large is meant a firm with substantial plant, hundreds of employees, occupying a considerable area, and with an important position in the particular trade. Some of these are larger than many public companies. The medium firms fall into neither of these categories.

TABLE I

NUMBER OF FIRMS IN EACH SIZE GROUP

	Small	Medium	Large	Total
Textiles and Clothing	4	13	6	23
Engineering and Construction ...	8	8	7	23
Miscellaneous ...	9	14	3	26
Total	21	35	16	72

To give the age of the firm might disclose the firm's identity, so the following classification was adopted:

Group I. Firms founded in the post-war period (1946 onward).

Group II. Firms founded in the inter-war years (1919-1939).

Group III. Firms founded pre-First World War period (1900-1914).

Group IV. Firms founded in the nineteenth century or before.

Firms in Groups I and II are generally managed or owned by the original founders, in Group III by the second generation. In Group IV the experience is very diverse.

TABLE II

NUMBER OF FIRMS IN EACH AGE GROUP

Group	I	II	III	IV	Total
Textiles and Clothing... ...	4	7	5	7	23
Engineering and Construction	9	5	2	7	23
Miscellaneous	8	6	3	9	26
Total	21	18	10	23	72

If the figures in the two tables are brought together we arrive at the following results :

TABLE III

COMPARISON OF AGE AND SIZE OF FIRMS

	Age Group I	II	III	IV	Total
Size :					
Small	16	2	2	1	21
Medium	5	12	7	11	35
Large	0	4	1	11	16
Total	21	18	10	23	72

There is evidence here that the size of the firm is a function of its age. From a knowledge of the history of the individual firms it may be stated that most of them started as small firms. The post-war firms have not yet had time to grow large, although five of them have expanded their turnover and productive capacity from small beginnings to a medium size. Of the firms established in the inter-war period two only remain small and one of these is in a declining trade. Twelve have grown to a substantial size and four of them are large firms. Two only of the firms established before 1914 remain small, seven are medium size and there is one large business. In the oldest group one remains small ; although it maintains its position in the trade, it has difficulty in expanding. The remainder are equally divided between large and medium.

In general there has been a growth in the size of the firm, and as far as the present sample is concerned, conditions appear to

have changed little since Marshall compared industrial growth with the trees in the forest. (This inquiry concerns those businesses that still exist at the moment, and the present survivors give no indication of the mortality rate at different times).

Ownership and Management.

Table IV shows the ownership of the different firms, according to size.

TABLE IV

NUMBER OF SHAREHOLDERS

	2	3	4	5	6	7	8	9	10–20	21–50	Not known	Other	Total
Small ...	9	4	2	3	1	1	1						21
Medium ...	5	6	8	5	2	1			2	1	3	2	35
Large ...					1			1	4	5	2	3	16
Total	14	10	10	8	4	2	1	1	6	6	5	5	72

In general an increase in size is accompanied by an increase in the number of shareholders. The new small firm is often founded by one man, and his wife becomes a shareholder to satisfy the requirements of the Companies Act. Often she is also a director of the firm. Even where other people are concerned, brothers or business associates, the main driving force is often the main shareholder and managing director. Sometimes a number of people combine together to start a business. Most of them have other business interests, or they are members of the family providing the capital; but they take little active part in the management.

As the firms increase in size, other people are brought in; sons join their fathers, brothers leave their paid employment to join the successful family business, and business associates, either suppliers or customers, often join the board. In the larger firm, which is in the second or third generation of ownership, the increase in shareholders is often caused by the sub-division of family holdings through inheritance. In a number of cases the main body of shareholders are on the distaff side, although the active management may be in the hands of the male descendants.

In some cases the active management is provided by non-family working directors.

In the case of the firms "not known," the organisation was described as "a family business." The five unclassified include three which are subsidiaries of other private companies and two which are legally public companies, although they are public only in the sense that the number of shareholders, for historical reasons, exceeds fifty ; they are not public in the sense of appealing to the general public for capital.

The actual controlling interest is much closer than the number of shareholders. The extent to which the working directors between them own the capital of the company can be seen from the following table.

TABLE V

OWNERSHIP AND CONTROL

	One	Two	Three	Four	No Control	Total
Small	10	7	2		2	21
Medium	9	14		2	10	35
Large...	1	3			12	16
Total	20	24	2	2	24	72

Controlling interest has been defined as owning over fifty per cent of the shares. It can be seen that in 44 out of 72 firms the real ownership and active direction is in the hands of one or two people. In effect they are individual businesses of the classical model, except that they have taken advantage of the Companies Acts to incorporate their businesses. The older and larger firms shade almost imperceptibly into the public company form of organisation where no individual, or group of individuals, holds a controlling interest.

The Supply of New Capital.

Of the sample of firms, 21 were started either at the end of the war or in the post-war period. The original capital put in by the founders was as follows :

TABLE VI

ORIGINAL CAPITAL OF POST-WAR FIRMS

Under £1,000	8 firms
£1,001 to £2,000	3 „
£2,001 to £3,000	2 „
£3,001 to £5,000	2 „
Over £5,000	4 „
Not known	2 „
	21

Of the eight firms started since the war with less than £1,000 capital, five have been engineering. Although the industry as a whole may be regarded as highly capitalised using a great deal of expensive equipment, it is divided into many branches and stages of development. There are many small and moderate size firms which operate in various stages of production, supplying component parts of finished products, or doing operational work on parts, such as machining rough castings. While the large firm may perform all the necessary operations within its own plant, each task is done in a separate department. What is done in a separate department of a large firm may be performed in another firm which is a separate financial entity. The large firm has to decide whether it will make the components itself or buy from an outside firm. In post-war demand conditions the large firms have often been pressed to fulfil their orders from customers out of their own resources, and have been glad to put out part of the work to other firms. This has given the opportunity for the smaller firms to get established.

The possession of even a few machines will enable the owner to contract for certain work. As the capital increases the number of machines and employees can be gradually increased. Generally such firms have been established by people who are skilled engineers ; they know their machines and can adapt them to suit certain purposes and they can buy second-hand machines to make a start, so that the capital requirements can be moderate.

Even in textiles, where the machine costs are higher, a limited capital may be employed by concentrating on part only of the total process, by buying in partially completed goods for finishing or sending out goods to other firms which possess the equipment for finishing. In some cases there has been no cost of materials, as

work has been done for other firms on their partly-made goods on a commission basis. Sometimes the firm starts as an offshoot of another business and one or more of the directors continues his own business while providing part of the capital and management for a friend or business associate starting a similar business.

The sources of capital have been very varied. Sometimes it was money saved while working in the same industry, or money accumulated during the war, mortgages on house property, and loans from members of the family. Often it was the combined effort of two or three people in similar circumstances, each putting up a comparatively small sum. What has not been done is to rely on capital from "outsiders," although this has often been offered. The man in a new business fears that he may lose control if a "financier" is brought in, and that the latter will be more interested in a quick profit than in the permanent building up of the firm.[1]

Pre-war Companies.

Eighteen of the firms in the sample were established during the inter-war years, with the following initial capital.

TABLE VII

ORIGINAL CAPITAL OF INTER-WAR FIRMS

Under £1,000	2 firms
£1,001 to £2,000	1 „
£2,001 to £3,000	2 „
£3,001 to £5,000	2 „
Over £5,000	7 „
Not known	4 „
	18

While the older firms in general had a larger starting capital, even without allowing for the lower pre-war level of prices, too much should not be read into this. The post-war firms are heavily weighted with the small engineering firm ; the inter-war firms consist of 7 textile, 5 engineering and 6 miscellaneous. The inter-war firms generally were started on a more complete scale and so required a larger investment in equipment. Some represented transfers of capital from a declining industry into one

[1]See also F. R. Jervis, "The Supply of Capital to New Businesses," *Journal of Industrial Economics*, February, 1956.

which offered better prospects, thus showing a mobility of resources with changing demand. In spite of small original capital, many of these firms are now substantial businesses.

No useful purpose would be served by analysing the original capital of the older firms in Groups III and IV. The main interest lies in the capital supply under post-war conditions. But it is known that most of these earlier firms also started with a very small capital and have grown from small beginnings.

Outside Capital.

In addition to the capital resources of the entrepreneurs other sources of capital are available from outside. For medium term needs one source is the purchase of equipment by hire-purchase. The following table shows the use of hire purchase by firms, according to age group.

TABLE VIII

NUMBER OF FIRMS USING HIRE PURCHASE

Age Group	Total Number of Firms	Number Using Hire Purchase
I	21	8
II	18	1
III	10	0
IV	23	1
Total	72	10

Hire purchase facilities are used by the new firm that is trying to get established and must obtain its first installation of machinery in large and indivisible units. Once the firm is established it may be some time before further installations are required. If it is successful it will have acquired other resources out of which it will purchase new or replacement machinery.

One firm reported that hire purchase had been used because the difference between the rate of interest charged and that charged by the bank was comparatively small, and in any case was a taxable expense, while the immediate supply of the machinery was very helpful. "The machines pay for themselves in the amount of trade they bring in. It is advisable to build up goodwill with the finance company. The first transaction is subject to considerable inquiries and references, but if payments are made

promptly, subsequent contracts are made with the minimum of fuss and are almost automatic".[1] Another firm obtained the bulk of its machinery on hire purchase by making out banker's drafts for the payments and there were no further formalities.

An engineer who had started with little capital used second-hand machinery for the first three years, then he bought new machinery on hire purchase. "The work that it did paid for the men's time, the charges on the machinery, and left a profit, whereas the old machinery paid the men's time with very little profit left over."

Another engineer who had started with £400 capital and £400 on loan, installed £1,200 of machinery bought on hire purchase. The firm was very profitable and the debt was paid off before the due date and new machinery bought out of accumulated profits. Another started with £1,500 as working capital but added £500 worth of machinery on hire purchase.

The knitting trade requires a larger capital investment in expensive machinery and although one group of people were able to raise over £10,000 between them, this was inadequate. The machinery was, therefore, bought on hire purchase and "no dissatisfaction is felt at the cost of this method of financing. The machines are earning their cost because of the high volume of profitable sales, and the interest charges are a taxable expense."

One firm which, after initial difficulties, has expanded rapidly out of profits chose the hire purchase method of finance as a policy rather than out of necessity. "We have preferred to keep our assets liquid so as to take advantage of favourable opportunities. Some good bargains have been obtained at auctions because we had the necessary liquid resources."

With the older firm the replacement of fixed assets will tend to follow a fairly regular pattern so that there is not the immediate financial strain as falls on the new firm. The one old firm that has used hire purchase reported that it once used it for the purchase of lorries but "it is more trouble than it is worth."

[1]All statements in inverted commas are the comments and opinions of the directors and officials of the firms interviewed. Whether the opinions are "right" or "wrong" from the academic point of view is immaterial in this connection. They are the basis on which decisions are taken. Obviously no direct reference can be made to the sources.

While it might appear from the above that hire purchase plays a small part in business finance, that part is still important. It enables the small firm to get established and to accumulate the financial strength to be able to dispense later with such methods. On the whole the firms interviewed expressed satisfaction with this method of finance. Although the interest rates are high, the earnings on the acquired assets have justified the cost. One firm bought very expensive equipment in this way; the cost was justified as they had more work than they could manage.

Yet there were firms in a similar position that did not buy on hire purchase at all. Some of the businessmen had a "moral" objection to incurring the debt and valued their complete financial independence at the cost of a slower expansion.

Bank Advances.

The Banks are the traditional source of outside working capital. In the following investigation a distinction has been drawn between those firms that normally use bank advances, those that make no use at all, and the intermediate group which has at some time or another during the post-war period made irregular use of bank advances.

TABLE IX

USE OF BANK ADVANCES BY AGE OF FIRM

Age Group	I	II	III	IV	Total
Use banks	13	7	4	13	37
No use	7	6	5	8	26
Irregular use	1	5	1	2	9
Total	21	18	10	23	72

TABLE X

USE OF BANK ADVANCES, BY SIZE OF FIRM

Size	Small	Medium	Large	Total
Use banks	13	18	6	37
No use	7	11	8	26
Irregular use	1	6	2	9
Total	21	35	16	72

New firms placed considerable reliance on the banks for their working capital, and so did the small firms. Firms founded in the inter-war period made less use ; on the whole they have built up their resources and have accumulated sufficient reserves out of profits to be more self-supporting. There is less reliance in Group III, for the same reason. It is when we get to the oldest firms that we find a greater use of bank credit. Their capital position has become stabilised, and often a bank advance of varying amounts is a permanent feature of the balance sheet.

In the majority of cases the borrowers expressed satisfaction with the services rendered by the banks.[1] Directors stated that though they are small or medium size, and are private companies, this has not handicapped them in their dealings with the banks. Even new, small companies have been helped over their initial difficulties although there has been little tangible security to offer. When any loan is requested a good case has to be made out and if this can be done the private company is at no disadvantage compared with the large public company, bearing in mind that its requirements are less.

One founder borrowed £500 from his bank on the security of his house. In another case when the firm was started as a partnership unsecured loans were given, though when the business was converted into a company, securities had to be deposited to cover the overdraft. "The bank has treated us very well." In a number of cases part of the original capital went into property used for the business of the firm, or the owners had house property, so that there was something to pledge as security. Yet this did not set the upper limit to bank finance ; "On occasions the overdraft exceeded the limit set by the bank without any difficulty, and banks are very helpful when they have confidence in the person." Without tangible security, another firm obtained overdraft facilities up to the limit of the outstanding debtors. Others secured overdrafts on the guarantee of one or other of the directors. One director gave the impression that he was welcomed with open arms by the bank manager because eighty per cent. of the output was exported, most of it to dollar markets.

A founder of a small engineering firm who preferred bank advances to hire purchase for new machinery as it was cheaper

[1]This investigation took place, of course, before the "credit squeeze."

and more flexible stated that "bank advances are ideal if you can get the right terms." The bank manager has a knowledge of the past record of the firm and in particular of the proportion of profit drawn out by the owner and so can assess the risk of a loan. In the case of a new business, which *ipso facto* has no profit record, and where the founders have put all their assets into the business, bank managers have stated that "each case is treated on its merits" and advances are made on the balance sheet alone, or on an assessment of the personal qualities of the borrowers.

With the older firm with a good steady record there are accumulated assets to pledge as security. One firm reported that there is a floating charge on the assets to cover an overdraft where needed. "This is obtained without any difficulty." On contract work there may be delays at the site, or there may be hold-ups at the ports, and when these difficulties occur "the bank can look after the old established private company." In many firms the overdraft facilities considerably exceed the paid up capital, and in contract work fluctuate with the volume of outstanding contracts. One firm undertook a rapid expansion in its fixed assets out of a bank loan, and this was later supplemented by a loan from the I.C.F.C. Just after the war one business man with a capital of about £2,000 plus the premises occupied, wished to take advantage of the opportunities which presented themselves in his trade. He was advised by a solicitor to put his case to the local manager of a bank (not his own) and to ask for a loan of £10,000. After examining the case the manager said "What you really need for this development is £20,000," and overdraft facilities for this amount were granted on personal recommendation alone. The working directors of this firm have since limited their drawings from the profits so that the overdraft has gradually been reduced. Another firm in difficulties found that by changing its bank it could get better facilities, in fact "it is possible to play one bank against another."[1]

[1]Bank managers have confirmed that there is no hard and fast rule with regard to the making of loans. Generally the larger the branch the greater the amount of discretion allowed to the local manager, even if he has to get subsequent confirmation from his head or area office. Partly it depends on the initiative and the amount of responsibility that the bank manager is willing to assume. As one banker put it, "There are managers and managers !"

Even where firms have not been able to get all they asked for it was admitted that this was "probably because of our profit record" or "because of the past history of the industry."[1]

The twenty-six firms that made no use of bank advances were strong enough financially to dispense with this method of finance. Some were very strong indeed with considerable bank deposits built up out of profits. These are retained to cover any eventualities. It was curious to find many examples of business men who had the same objection to borrowing from a bank as they had to hire purchase contracts. They preferred to expand slowly, consolidating each step, rather than to rush ahead and to risk other people's money.

The irregular users have no objection to borrowing, but their needs are not continuous. Some find their cash resources running down when tax payments are due, so a temporary advance is arranged. Others faced with steady production but seasonal sales have an overdraft at one period of the year but large balances when payments come in from their customers. Peculiarly, fear of shortage of materials has caused overdrafts. During the post-war material shortages many firms bought and "hoarded" raw materials whenever they were available. If this caused a financial strain money was borrowed from the bank temporarily. Later when the material position was easier firms could work with a smaller amount of stock and work in progress and bank loans were unnecessary.

Business expansion is often discontinuous. Even where a firm is following the principle of expanding out of profits it may require the installation of an expensive indivisible factor. Many firms have obtained temporary advances, repayable over a period of years, to cover some extensions.[2] These are gradually paid off

[1] In some cases where dissatisfaction was expressed with the magnitude of bank accommodation offered, examination of the balance sheets showed that the banks were justified in being cautious. Loans are not appropriate where financial stringency is caused by lack of profits, or even by excessive withdrawals of profits.

[2] While advances are in theory repayable on demand, and while the normal advance may be for a period of about six months, banks are often in practice prepared to make longer advances. "There is, in fact, a lot of medium term and even long term lending by banks while masquerading as short term lending" (Sir Oscar Hobson in *The Banker*, February, 1956).

so that the firms normally revert to a credit balance at the bank. Sometimes expansion means a change in the methods of distribution, so that instead of dealing through wholesalers capital is tied up in stock and debtors through dealing direct with the retail trade. This working capital has likewise been provided by the banks.

CAPITAL FROM CREDITORS.

Another source of outside capital is obtaining goods from suppliers on long term credit. In a few cases new firms have made special arrangements with their suppliers to defer payments until a later date ; and suppliers have been willing to do this to establish a market. But on the whole the private company is in the same position as the public company in its relations with its creditors. The terms of credit seem to depend on the custom of the trade and not on the age, size or standing of the firm. The terms received from suppliers usually bear a relationship to the terms offered to their own customers, and the private company in any trade has little option in the matter and conforms to the "custom of the merchants."

In light engineering contracting supplies are normally paid monthly in arrears so that as payment is made to the firm for work done the money is available to pay for materials used. In the textile trades the terms of payment have been tightened up since the war so that the money has to be found quickly. This can be done where the firm sells to other sections of the trade where prompt payments are made, or to wholesalers who are organised to effect prompt payments. If the firm deals direct with retailers it will require a larger investment in working capital to cover the slower settlement of accounts, which is usually on a monthly basis.

In heavy engineering and contracting part of the finance is provided by the customer. Monthly progress payments are made against work done, less a retention of ten per cent. until the completion of the contract. It is possible for a well organised contractor to have an annual turnover of fifteen to twenty times the amount of his own capital. Merchanting firms are in the weakest position, as they normally pay monthly but take much longer to get in their own accounts, especially in building.

Many firms complained that while they themselves paid promptly they had slow paying accounts among their largest customers, especially the large companies, public authorities and government departments. The large company too often seemed to have a slow and complicated accounting system that caused delay in settlement. Government departments took up to six months to pay but the firms always allowed for this in tendering for government contracts. One successful firm had refused to supply government departments because of the difficulty in getting payment.

On the whole credit is "neutral" in the case of the firms in this survey ; where a firm expands the increase in creditors is matched by the increase in money outstanding from customers, and the private company is in no better or worse position than the public company.

Capital from Profits.

The traditional source of business expansion in the nineteenth century was the accumulation of profits in the firm. The extent to which expansion out of retained profits can be achieved will depend on the profitability of the business ; on the willingness of the entrepreneur to forego present consumption in the interests of future expansion ; and (of greater force in recent years) on the balance available after taxation has been paid. The first and the last are related. A period of high profitability even with high taxation may well leave a greater surplus for expansion than a period when taxation is low, but business is depressed.

As was to be expected, the founders of new businesses, generally with a small capital, had the ambition to expand into something larger. Partly this was due to the increased financial reward to be expected from a larger business, but just as important was the motive of building up for the sense of achievement that this brings. The owner-manager with a given income saves more than an employee with the same income. While the motive has been there, so have the opportunities. In most industries demand has been good and the period of rising prices has provided the setting in which the business firm has the confidence and the resources to look to the future. Where profits have been good, taxation has been paid (in arrears), drawings have been limited

and the balance has gone into accumulation. Up to a point this accumulation is at compound interest.

Some post-war case histories will illustrate this. One small family engineering business pays the owners the same salary as they would earn as employees, the rest being left to accumulate. It is now well established with a new product in the export markets. One man started with £1,200 after the war and in six years the company's reserves exceeded £5,000 : all net profit after meeting the managing director's salary was left in. Another man started with £500 and now has a capital of £3,000 ; and another went from £800 to over £6,000 in the same way. Six people working together started with £700 plus a large loan from one of the directors. In six years the capital had increased to £15,000 and the loan had been paid off. "There have been no extravagant drawings and the directors are content to use small horse-powered cars."[1]

In the group of firms founded during the inter-war period there were many in a similar position. They were still in the first generation of management and there was still the enthusiasm to build. While taxation was a greater burden than before the war, on the whole profits were easier, and "even half the profits is better than none at all." Where firms in this group were not expanding, the cause lay in the lack of profits due to contraction or changes in demand in those industries. One large firm had been incurring losses during the slump years but since the war good profits have been earned. The firm has grown five-fold in this period. The amounts placed to reserve are about three times the amounts distributed. "The expansion has been possible because the shareholders do not bleed the company." Another businessman built up his capital from £1,000 to £10,000 because he was prepared to live moderately.

One firm has been handicapped by its own success. The directors would be willing to leave larger amounts in the business but are unable to do so. Some of the shareholders are in the surtax class, and the Inspector of Taxes insists that a reasonable amount of the profits be counted as dividend ; otherwise an individual assessment would be made on the surtax shareholders

[1]See *Journal of Industrial Economics*, February, 1956.

and the amount charged to the company. As a result the post-war expansion has had to be financed from various outside sources.

In the older and larger firms founded before 1914 there is less tendency to expand. The profits are there, but instead of being used to extend the productive assets they are used to increase the liquidity of resources so as to be independent of the banks or even to accumulate cash balances. Many of the owners remember the inter-war depressed conditions and so do not want to over-reach themselves in the easier conditions of today.

Moreover the desire to expand may have diminished. The newer, smaller firm can expand until the owners reach a position of comparative affluence. There is much less incentive after this position has been reached. The tax burden becomes progressively greater and many directors have pointed out that an increase in profit would make no significant difference to their standard of living after tax. The well established profitable business tends to be more static.[1] The marginal significance of extra income after tax does not compensate for the additional work that would be required, nor for the additional risk incurred in physical expansion of assets and markets if conditions should change.

In contrast to the medium or large sized firm that is still expanding out of its own resources and approaching the point where it may eventually be converted into a public company, one often finds a firm in similar financial circumstances, where for a variety of reasons the directors are content with their present size. Often it is because they believe that they are at the optimum; they value the closer relationships of a reasonable size and are men of moderate ambition.

SUMMARY.

The inquiry was concerned with the supply of the factor of capital. While it is true that scarcities of capital were discovered

[1]This is partly due to the diffusion of ownership of the share capital (see Table IV above). In the newer and smaller firms the profits tend to go as *salaries* to the working directors (subject to earned income relief) and the balance not used for living expenses is left in on current account. Where the ownership becomes spread over the descendants of the founders, especially on the female side, the firm has to pay dividends on *profits* (no earned income relief), otherwise the beneficial owners would get no return from their holdings. Moreover the tax burden is proportionately heavier.

—otherwise capital would not be a factor of production—there was no evidence to show that the private company was experiencing any special difficulty in obtaining the capital required. In post-war conditions other factors were often scarcer and prevented industrial expansion even where the capital was available. Firms with the money had difficulties over "land" and "labour" supply. There is no hard and fast division between the private and the public company. From an economic point of view the situation of the larger private company shades almost imperceptibly into that of the smaller public company.

The experience of the firms surveyed has been very varied. Some have expanded using almost every method of obtaining finance while others have relied on their own resources only. Some firms were kind enough to allow access to their balance sheets, and the following details illustrate some of the diversity of experience.

1. This engineering firm in four years increased its assets from £16,500 to £38,000 from the following sources:

		(£000)		(£000)
Capital	from	6	to	13
Reserves	„	2	„	6
Creditors	„	4·5	„	3
Overdraft	„	4	„	16
		16·5		38

It is expanding faster than it is able to accumulate profits. The increase in the issued capital has come from the owners, and their friends who are willing to venture a limited sum and to wait for a return on the investment. The extent of the bank's participation can be plainly seen. The money has gone in extending the premises and in building up stock.

. . . .

2. In six years since the war another engineering firm increased its assets from £6,400 to £78,100 from the following sources:

Increase in Assets		(£000)	Financed by increase in		(£000)
Buildings	from	0 to 6·8	Capital	from	0·7 to 15·0
Machinery	„	3·5 „ 8·6	Loan	„	3·0 „ 0·0
Stock	„	0·7 „ 13·4	Reserves	„	0·0 „ 50·0
Debtors	„	2·2 „ 18·3	Mortgage	„	0·0 „ 5·2
Cash	„	0 „ 31·0	Creditors	„	1·5 „ 7·9
			Overdraft	„	1·2 „ 0·0
		6·4 78·1			6·4 78·1

The firm started in rented premises which it soon outgrew and took over the present factory on mortgage. The factory is now too small but a second change would present organisational problems which the directors do not feel prepared to face at the moment. The increase in the capital, the repayment of the loan, and the reserve fund, have all come out of profits. The overdraft of the first year has been replaced by a £31,000 cash balance.

3. On a smaller scale, the firm that started with £400 in cash and £400 on loan and a hire purchase debt of £1,200 (but with a guarantee of any bank overdraft required) has increased its assets from £2,000 to £15,000 in four years. Creditors total £5,000, and capital and reserves accumulated out of profits are £10,000.

These figures do not tell the whole story as money was invested in the firm by an outsider, and as there was a difference of opinion he had to be bought out.

4. A rapid expansion was achieved in two years by a firm in the textile trade by using every method of finance. The increases are shown by the following figures (£000) :

	Year 1	Year 2		Year 1	Year 2
Machinery	12·7	17·0	Capital	10·0	14·8
Stocks	6·6	16·0	Reserves	0·0	8·6
Debtors	1·7	9·0	Hire Purchase ...	6·0	7·3
			Creditors	2·0	7·3
			Overdraft	3·0	4·0
	21·0	42·0		21·0	42·0

The increase in the capital has come from additional amounts being invested by some of the directors. There has been no difficulty in obtaining capital ; various business associates have wanted to join the company but have been refused as the founders did not want to lose control. Premises are rented. (In fact the landlord took the favourable opportunity of becoming a shareholder!).

5. Another textile firm was started before the war with a capital of £100. The founder was already in an allied branch of the trade so all profits of the new business were left there and added to capital, which has increased from £100 to £17,000. The firm is

"hoarding" £10,000 of cash which it cannot use as the firm has reached the limit of physical expansion in its present building.

6. Good profits do not necessarily mean an easy financial position. Much of the strain of the higher price level may be thrown on the large merchanting firm which is fulfilling its traditional role of financing stocks for its customers. The following figures illustrate the changes in the balance sheet over four years of a large firm of merchants and wholesalers :

ASSETS (£000)

Year	1	2	3	4
Fixed assets	25	25	30	34
Stocks	37	44	55	50
Debtors	119	140	159	200
Total	181	209	244	284

LIABILITIES (£000)

Year	1	2	3	4
Capital	10	10	10	10
Reserves	27	31	40	47
Profit balance	10	16	10	21
Tax reserves	9	10	12	15
Creditors	93	103	129	157
Bank overdraft	32	39	43	34
Total	181	209	244	284

While the reserves and profit balances have increased by £30,000 in the four years the firm has to rely heavily on its suppliers and the bank to keep going. The root cause of the trouble is that the firm's customers are by nature and tradition very slow payers. This has meant that the great increase in trade has brought no increase in the incomes of the working directors as payments for tax and the amounts placed to reserve both exceed the total of the directors' salaries. It is this situation which is partly the cause of the present financial stringency. The directors know that something should be done, to re-organise the firm's trade, to cut out

the poor payers and to make some arrangements for additional permanent capital; but with little marginal financial reward after tax, nothing is done.

7. Even in an old established branch of the textile industry, where little progress would be expected, one firm has made a three-fold increase in its assets in seven years. The sources of this capital are:

		(£000)	(£000)
Increase in capital	from	10	to 20
„ „ reserves	„	8	„ 53
„ „ creditors	„	14	„ 30
		32	103

The turnover has increased from £70,000 a year to over £400,000 and has been financed internally. The increase in capital was a transfer from reserves. The managing director draws a regular salary and the rest of the profit is left in on current account.

Many other firms present a similar experience to the above examples. In the majority of firms finance was not a major cause of difficulty. It was often scarce, and the problem had to be solved in one or other of the various ways illustrated above. But very often other problems were more pressing. For example, expanding faster than the existing accommodation, scarcity of suitable labour, and in a few cases, the contraction of markets. But on the whole, post-war demand has provided the setting in which companies can make profits, and in the majority of cases a large proportion of these profits have been ploughed back into the business. In addition other sources of capital have been used freely; increases in creditors, bank advances where necessary, and long term loans. Capital from friends and business associates has been used but the one source which is viewed with the gravest suspicion is the offer of finance from strangers. Directors of family businesses generally view them with affection, and fear that control may pass to other hands if outsiders are admitted. Even if control is retained the outsider's interests are purely financial, whereas the family business is something more than a dividend-paying machine.

Combined figures from eleven firms in different trades of all ages show an increase in capital of over half-a-million pounds :

	(£000)
Increase in fixed assets, machinery etc.	105·4
„ „ stocks and work in progress	198·6
„ „ debtors and pre-payments	249·4
„ „ cash resources	43·9
Total	597·3

This has been financed by :	
New capital subscribed	17·0
Reserves ploughed back out of profits	323·1
Increase in creditors	243·0
Bank overdrafts	14·2
Total	597·3

The increase in creditors is more than matched by an increase in debtors, and cash resources have increased more than bank borrowings. (For industry as a whole the post-war increase in bank deposits considerably exceeds the increase in bank advances). The firms in the sample are not running into indebtedness *vis-a-vis* the rest of the economy. Little new capital has been put in. The exceptional cases are where firms are expanding faster than their rate of profit accumulation, and even there the bulk of the capital for expansion has been self-generated from the firms' own resources.

F. R. Jervis

Derby and District College of Technology.

Books Received

International Labour Office : *Production and Employment in the Metal Trades.* The Problem of Regularisation. Geneva. 7s. 6d., $1.25. 121 pp.

International Labour Office : *Social Aspects of European Economic Co-operation.* Report by a Group of Experts. Geneva. $1.50 ; 9s. 179 pp.

ROOSA, Robert V. : *Federal Reserve Operations in The Money and Government Securities Markets.* Federal Reserve Bank of New York. 108 pp.

WIGHTMAN, David : *Economic Co-operation in Europe.* Carnegie Endowment for International Peace, European Center. Prizes in International Organization. Stevens & Sons Ltd., Wm. Heinemann Ltd. 21s. net. 288 pp.

SHARMA, Dr. S. L. : *Some Trends of Capitalist Concentration in India.* P. C. Dwadash Shreni & Co. Ltd., Aligarh. 7/4/-. 218 pp.

I Problemi Dello Sviluppo Economico. Contributi Dell'Istituto di Scienze Economiche. Serie IV. Volume LIV. Societa Editrice, Vita E Pensiero Milano. 430 pp.

EISERMANN, Dr. Gottfried : *Die Grundlagen des Historismus in der deutschen Nationalokonomie.* Ferdinand Enke Verlag Stuttgart. 28 DM. 249 pp.

MILNE, Alastair, M. & LAING, Austen : *The Obligation to Carry.* Institute of Transport. An Examination of the development of the obligations owed to the public by British Railways. 5s. 129 pp.

MISRA, B. R. : *V for Vinoba.* The Economics of the Bhoodan Movement. Orient Longmans, Bombay. Rs. 3. 67 pp.

MUMMERY, A. F., & HOBSON, J. A. : *The Physiology of Industry.* Reprints of Economic Classics. Kelly & Millman, Inc., New York. $5.00. 215 pp.

SOVANI, N. V., APTE, D. P. & PENDSE, R. G. : *Poona : A Re-Survey.* The Changing Pattern of Employment and Earnings. Gokhale Institute of Politics and Economics, Publications No. 34. Rs. 15 or 25s. or $3.50. 555 pp.

HALMOS, Paul : *Towards a Measure of Man.* The Frontiers of Normal Adjustment. Routledge & Kegan Paul Ltd., London. 28s. 250 pp.

SZCZEPANIK, E. F. : *The Cost of Living in Hong Kong.* Hong Kong University Press. HK $5.00, abroad 7/6d. net ; US $1.20. 25 pp.

HARPER, F. A. : *Why Wages Rise.* The Foundation for Economic Education, Inc., New York. $1.50. 124 pp.

THE MANCHESTER SCHOOL OF ECONOMIC AND SOCIAL STUDIES

VOLUME XXV

1957

Reprinted by permission of the original publishers
KRAUS REPRINT LTD.
Vaduz
1964

Printed in Germany

Lessing-Druckerei — Wiesbaden

CONTENTS

No. 3.. September, 1957.

Page

THE CASH REWARD FOR AN HOUR'S WORK UNDER THREE PIECE-WORK INCENTIVE SCHEMES .. — Tom Lupton, Sheila Cunnison — 213

FACTOR ENDOWMENTS, INTERNATIONAL TRADE AND FACTOR PRICES .. .. — H. G. Johnson — 270

THE LEGAL MINIMUM WAGE, EMPLOYERS AND TRADE UNIONS: AN EXPERIMENT .. .. .. — Brian McCormick, H. A. Turner — 284

A NOTE ON SECULAR SWINGS AROUND ALTERNATIVE TRENDS — Rosemary Walker — 317

The Cash Reward for an Hour's Work under three Piecework Incentive Schemes[1]

In recent years there has been much criticism of economists' explanations of the process of wage determination. Doubts have also been expressed whether present procedures for relating cash reward to work done by industrial workers are entirely appropriate to modern industrial conditions.

Sociologists have been responsible for much of the criticism. Mrs. Barbara Wootton, for example, has argued that, by itself, economic analysis has added little to our understanding of the actual process by which wages and earnings are determined.[2] She argues for a shift in emphasis to take account of social and psychological factors. Other students of the movement of wages and earnings, notably H. A. Turner,[3] have drawn attention to the discrepancy between rates negotiated at industry level, and the workplace rates which emerge in the process of interpreting industry rates to fit the local context.

Part of the discrepancy which Turner and others have noted between industry and workshop rates is probably due to the increasing application of Work Study, and to the introduction of piecework incentive schemes based on time-study data. Work Study engineers claim that by careful and systematic study of the content of industrial operations, a relationship can be established between task performed and cash reward which is more accurate and objective than any which can be established by bargaining. Arguing along these lines, some advocates of a national wages policy have suggested the possibility of creating a wages fund, shares of which could be allocated by using modern techniques of Time-Study and

[1]The research on which this article is based was sponsored by the Department of Scientific and Industrial Research/Medical Research Council Joint Committee on Human Relations in Industry, and financed from Counterpart Funds derived from United States Economic Aid.

[2]Wootton, Barbara, *The Social Foundations of Wages Policy*. London: Allen and Unwin, (1955).

[3]Turner, H. A., "Wages : Industry Rates, Workplace Rates and the Wages-Drift." *The Manchester School*, Vol. 24, No. 2, May, 1956.

Job-Evaluation. So, in addition to academic criticism, the increased adoption of Work Study and Job-Evaluation itself implies criticism of present procedures for determining earnings.

We doubt whether all the claims made for objectivity and accuracy on behalf of any known system of earnings determination can be supported. Our reasons for doubt will soon become apparent. We shall argue that, under piecework, there are so many factors which affect the level of earnings that at workplace level the determination of earnings should be a matter very much for evaluation and argument. Since this is so, opportunity for argument should be provided. On this ground we shall plead for greater worker participation in the determination of workplace job rates. The criticism which is implied in the facts we shall present is therefore of a somewhat different kind from those we have mentioned above.

Although the facts we present, and our analysis of them, have generated these doubts and criticisms in our own minds, we stress that our material is limited to three case studies so that it is quite possible that our conclusions do not apply generally. It is also possible that others will draw different conclusions from the facts we have gathered.

We have collected figures for three piecework shops showing the earnings of the workers there for a period of roughly six months. We worked alongside the workers throughout the whole period when they were making these earnings, and we are therefore able to say something about the factors which determined the pay which the workers drew each week.[1] Although the workshops studied differed markedly in many ways, we detected certain similarities. We found that regardless of the method of ratefixing there were anomalies in the rates. These anomalies, when considered together with the method of allocating work to the workers, explain many of the fluctuations in earnings which our figures show. Interruptions in the flow of work to operatives due to market conditions or to shortcomings in production planning, or even to chance factors such as absenteeism and human error, are all shown to affect, in

[1] Lupton worked in Factories I and II, and Cunnison in Factory III. We express our appreciation of the help and encouragement given by managers, Trade Union officials and workers, in the course of our enquiries.

some cases seriously, the amount which finds its way into the pay packet each week. And there are also other factors which we shall discuss later. Although we had no accurate means to measure the effort expended by the workers we observed, our impression was that in all the shops the factors we have just mentioned intervened to distort the relationship of equality between effort and reward, which was the rationale of all the incentive schemes. For this reason each of the schemes had features which the workers considered unsatisfactory. We shall describe how attempts were made by the workers, with differing degrees of success, to control some of the factors determining the level of earnings.

Our analysis of the earnings figures is designed to throw light on the two problems mentioned above, namely the relevance or otherwise of existing methods of wage determination to the problems which are encountered from day to day in the workshops, and the validity of claims made for methods of wage payment designed to equate reward and task performed. The analysis may be regarded by some as an argument against piecework incentive schemes as such. We make no such claim, and would emphasise that to substantiate such a claim would require a comparison of piecework and timework shops operating in a similar economic and technological environment. We have not yet made such a study.

Each workshop will now be separately considered. We shall describe the workshop and then consider in general terms the factors which determine the level of earnings. We shall then analyse in detail the earnings of selected workers over a period of roughly six months, tracing wherever possible the exact causes of fluctuations, and the outcome of efforts by the workers to control the factors influencing the level of their earnings. We shall then present conclusions based on the three studies.

Factory I.

(1) *Product and method of production.*

Factory I manufactured a wide range of rubber products and heavy clothing. The investigation on which this section is based was carried out in one of the workshops which assembled

garments. This workshop employed some sixty persons, all except two of whom were women. Most of the women were of middle age. About a third of the working force in the shop were married women, and most of these had children of school age. There were a dozen or so young single women.

Production was organised on a unit basis, and there were three units in the shop, each normally comprising sixteen operatives. Each unit was separately responsible for the completed assembly of garments from cuttings supplied from another workshop. Within the units, tables, machines, and other equipment were laid out so as to facilitate a smooth flow of parts from one operative to another throughout the process of sub-assembly and assembly of the garment. There was no conveyor to regulate the pace of work, the parts being passed from operative to operative. Thus, the speed of work and the quantity of output were regulated by the efforts of the operatives and by the smoothness or otherwise of the flow of parts from operative to operative within the units.

Formally, all the operatives were classified as semi-skilled, but there were wide differences in skill and experience between those workers at the beginning of the assembly process in each unit who prepared the parts for assembly, and those who were engaged on more complicated assembly operations in the later stages of the work sequence.

The operatives were of two categories, *assemblers* and *machinists*. There were usually twelve assemblers and four machinists in each unit, although two of the assemblers did some machining as part of their work.

(2) *The incentive scheme.*

Each of the many operations required to assemble a garment had been time-studied, and after allowances of various kinds had been made, each operation was given a cash price. One of the elements in the price calculation was a skill allowance, designed to ensure that greater skill and experience were rewarded. The job prices were related to the negotiated minimum wage for women workers in such a way as to ensure that the average worker expending average effort could earn more than the minimum. The allowance for skill ensured that

the more highly skilled workers could earn a great deal more than the minimum. The incentive system was a straight piecework system, operatives being paid weekly according to the number of operations completed. The minimum wage operated as a "fall-back" rate, should the earnings of any operator fall below the minimum.

The system was simple for the operatives to understand. Each bundle of cuttings (the material for a complete garment) was delivered to the unit together with a sheet of dockets, each one of which represented an operation to be performed on the garment. The worker completing the operation would detach the appropriate docket from the sheet. All the dockets were handed over each evening to the chargehand, who then forwarded them to the wages office. Since the prices of the jobs were known to the operatives it was a simple matter for them to calculate their earnings each day.

(3) *Method of job allocation.*

Within the production units each operative was responsible for a small proportion of the large number of operations, and each operative specialised on certain operations. Thus, the co-ordination of the work of the operatives within the units was built into the layout, and was not formally the responsibility of any one operative or any group of operatives. After collecting work from someone else each operative completed the operations which were assigned to her and then her responsibility was at an end. Other operatives collected the work as she completed it and began the operations assigned to them. Thus the co-ordination of the work of the operatives was designed to be mechanical in its operation.

Provided that everything went according to plan, so that each operative was working on operations that she had been trained to perform, and was receiving a steady supply of parts ; and given that the job prices reflected equitably the effort expended at different levels of skill, no personal intervention by the shop supervisor or by anyone else was necessary. But if for any one of the many reasons to be mentioned later co-ordination broke down, the shop supervisor would intervene to transfer operatives from one job to another to maintain an

even flow of work. In situations like these allocation of work became a personal matter between the supervisor and the operatives concerned.

In this shop the role of the shop supervisor in the allocation of work was not a positive one of co-ordinating the work of the operatives in her charge, but the negative one of correcting faults in the impersonal system of allocation.

The system of job prices served a dual function. We have already shown that it was designed to offer an incentive to the operatives to extend their best efforts. In practice prices affected the co-ordination of the work of the operatives who, trying to maximise earnings, might produce too few parts to keep the person next in the flow of work continually occupied. And this problem was, of course, related to the problem of the size of the production unit, and to the division of labour within the unit.

With so many factors to consider, it is probable that when planning the layout and the method of job pricing, the management decided on a scheme which seemed to offer best results but which was bound to fall short of perfection in operation. If this were so, and the system of job-pricing were free from gross anomalies, it would be reasonable to expect that any marked variation in the level of earnings of operators from week to week would be traceable without difficulty to differences in effort expended.[1]

Observations in the workshop over nearly six months, and direct experience as an operative in one of the units, make it possible to say that there was little difference from day to day in the pace of work. Some operatives could work more speedily than others, and some were more skilful than others, but when each one is considered separately it can be said that they worked at a pace which was fast and even and which changed little from day to day. There are two minor exceptions to this general statement : the practice of "sweating" over short

[1]This is, of course, the rationale of any incentive scheme, and its fairness or otherwise is judged largely by this criterion.

periods,[1] and changes in pace due to uncertainty associated with changes in style of garment.

Any gross fluctuations in earnings are thus to be explained largely but not entirely by reference to factors other than differences in effort expended. We now go on to discuss these "factors influencing the capacity to earn." Then a number of examples will be given showing fluctuations in earnings and an attempt will be made to show the influence of the factors in each case.

(4) *Factors affecting the capacity to earn.*

(*i*) *Waiting time.*

In this workshop waiting time allowances were not negotiated "on the spot" to compensate for actual waiting time. Observations of the system of booking in work suggest that no accurate records were kept of time spent waiting for work.

(*ii*) *Variety of materials.*

Some cloths were much more difficult to handle than others. Some were difficult to keep clean, others induced eye strain, and so on. When calculating job prices, management attempted to compensate for these factors. Operatives often complained that the allowances were inadequate, and it will easily be appreciated that management found it extremely difficult to calculate allowances which accurately reflected the effect of changes in materials on the output of particular operatives. It is possible therefore that in some cases difficulties were overestimated, and in others underestimated. Thus variety of materials must be considered as a factor influencing stability of earnings.

(*iii*) *Variety of designs and length of run.*

Changes in the design of garments coming into production were frequent, and the length of run on any one design varied. For example, a unit might run for two weeks on one

[1]The operatives themselves distinguished two paces of work—"steaming" and "sweating." "Steaming" was defined as a fast but not a "killing" pace and was associated with situations of long runs of work at prices which promised good earnings. "Sweating" described a pace that was "killing" over short periods. "Sweating" was associated with conditions in which the price was "tight" and in which some sacrifice of quality was made to prevent a slump in earnings.

design then change and run for six weeks on a design which differed radically. Since different designs presented different problems of assembly, a change in design often affected the output of an operative while speed was developed on a new sequence of movements. The impact of a new design acted differentially. For some operatives a particular design would present few problems of adjustment, for others it would raise many difficulties. When adjustments were being made to a new style mistakes in assembly were sometimes made. When these were detected by the inspectors or by other operatives time was spent correcting them for which no specific allowance was made in the job prices. Thus although an operative might expend great effort when a new style came into production, earnings might well be reduced during the period when familiarity was being gained with a new operation.[1]

(*iv*) *Allocation of Work.*

We have already said that the system of work allocation was designed to be impersonal and mechanical. The flow of work though the factory, which was planned by individual managers and supervisors, determined the work which an operative would be doing at any time. But when there were interruptions in the flow of work due to some imbalance in the planned production flow, or when the units were depleted through absenteeism, gluts of work accumulated in some parts of the units, while shortages developed in others. The shop supervisor would then transfer work from unit to unit, or transfer operatives from one operation to another, to keep the work flowing smoothly through the units. When operatives were transferred their earnings were sometimes affected while they were familiarising themselves with a different task. Since transfers did not affect all operatives at the same time, and since they did not occur regularly and systematically, it would have been difficult to have computed average "transfer" and "re-learning" allowances, and so far as we know the attempt was never made.

[1]Sometimes operatives complained that they were unable to make certain designs pay but could give no rational explanation. It was "just a bad coat to make money on."

Thus allocation of work tended to affect the stability of earnings. It must be said however that the chargehand's motive in making transfers was to offset the effects of waiting time on the earnings of operatives. Thus, as well as acting to keep the work flowing, the personal system of work allocation also acted to correct anomalies in the impersonal system, and must also be judged in this way.

So far, all the factors considered have been those which either reduced the earning capacity of the operatives or else led to day-to-day or week-to-week fluctuations. We now consider those factors which act to modify these effects.

(5) *Modifying factors.*

(*a*) *"Dead Horse."*

"Dead Horse" is the name given to work for which the worker has received payment, but which he has not yet completed. It is well known in this section of the garment industry as a technique for ironing out fluctuations. It has other effects, as will be shown in the analysis of Factory III. To offset anticipated fluctuations it was possible for operatives in Factory I to hand in dockets for work which they had not yet completed. In this way they could draw wages for work which they had not yet done. But the extent to which this practice could flourish was strictly limited. Dockets only began to circulate when work was delivered so it was relatively easy for supervisors to check work done against dockets handed in. Since the practice was frowned upon by management, and since "dead horse" had to be "worked off" in the long run, few operatives used the technique.

(*b*) *Debasement of quality.*

It was possible to compensate for a "tight" price by "sweating," i.e. by sacrificing quality for speed, but success in this direction was limited by the system of inspection, which was highly efficient in Factory I.

(*c*) *Voluntary absence.*

We recorded a small number of cases where slumps in average hourly earnings were avoided by voluntary absence. If it looked as if work next day promised to pay badly and work had been going well for a few days, the operative would take the day off, calculating that it would be worth while to spread payment for four good days over five. The resulting loss in weekly earnings would be compensated for by the leisure enjoyed, and average hourly earnings would remain higher than they would otherwise have been. Many operatives considered behaviour such as this to be immoral. It upset the flow of work and was a threat to the earnings of other operatives. There were so few examples of voluntary absence of this kind that the effects on earnings are difficult to trace. But the practice is worth mentioning as an attempt at control over the factors affecting earnings by some operatives.

We now present Tables and graphs of earnings of operatives in one of the units for a period of nineteen weeks. We shall attempt to account for some of the fluctuations in earnings by reference in each case to the factors we have outlined.

Table I shows how wide the range of variation is for some operatives. Operative 15 has a range of 1/8¾d. around an average of 2/5¼d. Operative 1 and operative 2 show small variations. So as to show that the ranges are not distorted by "freak" figures Fig. 1 has graphs showing average hourly earnings over the whole period for six of the operatives. Our field notes are, of necessity, not detailed enough to enable us to account for every variation in earnings for each of the operatives. We now consider each of these six operatives in turn to see how far it is possible to explain the causes of variations in earnings.

TABLE I

AVERAGE HOURLY EARNINGS (19 weeks)
FACTORY I

Operative	Category	Maximum hourly average	Minimum hourly average	Range	Average
1	Assembler	2/3½	1/11	4½d.	2/1
2	,,	2/7	2/0¼	6¾d.	2/4¼
3	,,	2/4½	1/8½	8d.	2/1
4	,,	2/6¼	1/9¼	9d.	2/4½
5	,,	2/2½	1/4½	10d.	1/8½
6	,,	2/11¾	2/1½	10¼d.	2/6½
7	Machinist	2/8¾	1/10½	10¼d.	2/6¼
8	,,	2/9½	1/10¾	10¾d.	2/3
9	,,	3/1¼	2/1	1/0¼	2/7¾
10	Assembler	3/8¾	2/7	1/1¾	3/2¼
11	,,	3/2¼	2/0¾	1/1½	2/6¼
12	Machinist	3/2½	2/0½	1/2	2/7½
13	Assembler	3/11	2/4	1/7	2/7¾
14	,,	3/11¾	2/4¼	1/7½	3/2
15	,,	3/1	1/4¼	1/8¾	2/5¼

Notes :

$$\text{Average Hourly Earnings} = \frac{\text{Weekly earnings (gross)}}{\text{Hours worked}}$$

$$\text{Average (col. 6)} = \frac{\text{Total of weekly averages}}{19}$$

Full earnings figures for those operatives underlined are shown in Fig. 1.

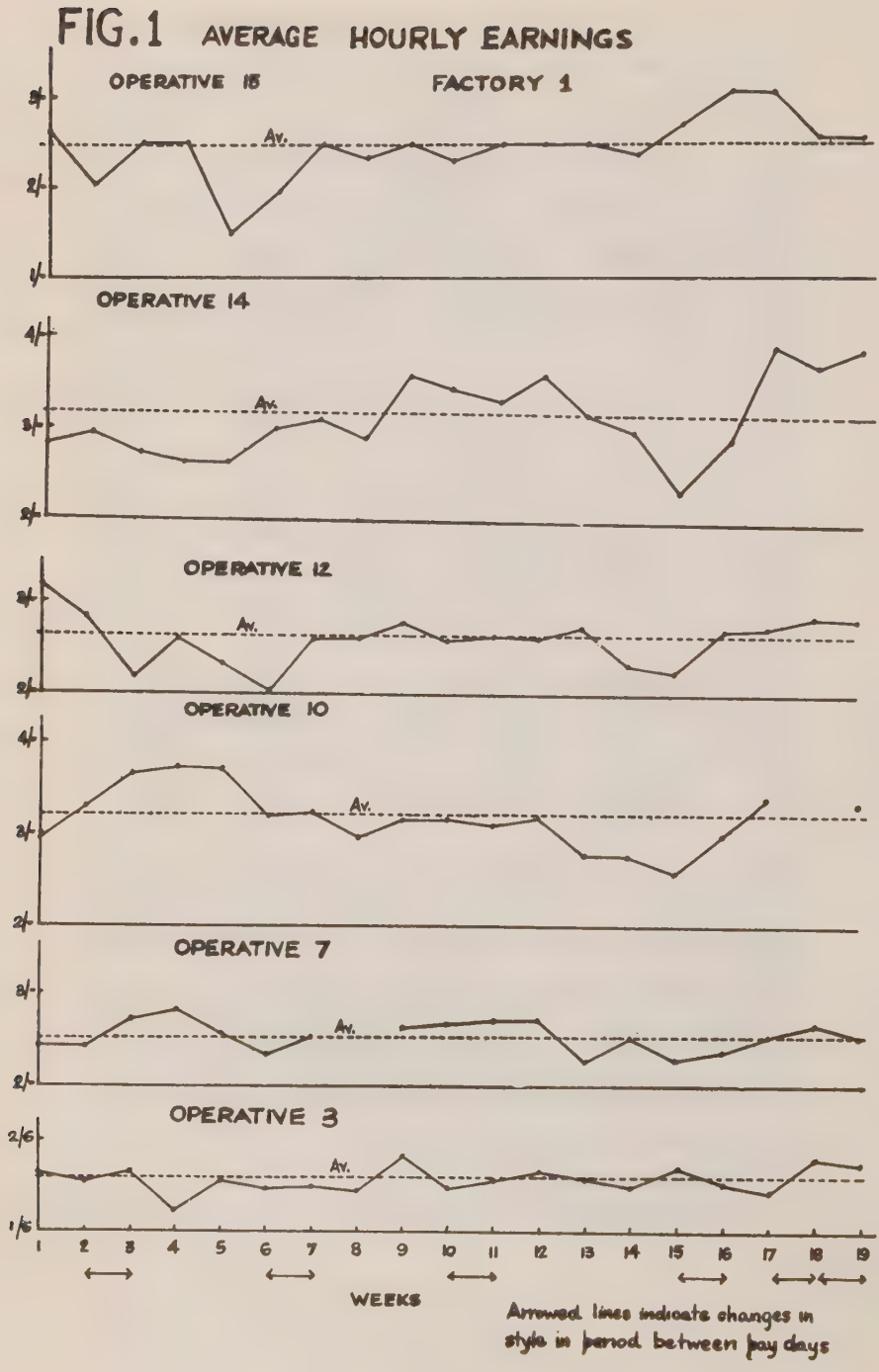


FIG. 1 AVERAGE HOURLY EARNINGS
OPERATIVE 15
FACTORY 1
3/-
2/-
1/-
Av.
OPERATIVE 14
4/-
3/-
2/-
Av.
OPERATIVE 12
3/-
2/-
Av.
OPERATIVE 10
4/-
3/-
2/-
Av.
OPERATIVE 7
3/-
2/-
Av.
OPERATIVE 3
2/6
1/6
Av.
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
WEEKS
Arrowed lines indicate changes in style in period between pay days

Operative No. 15.

Operative 15's earnings show a wide range of fluctuations and a lack of systematic variation. Fluctuations tended not to follow style changes very closely. Her earnings moved sharply up and down even in periods when styles were unchanged. The drastic slump in week 5 illustrates the effect that job changes and absenteeism could have on earnings. Another operative had been absent for a few days and Operative 15 was moved from her regular place in the unit to take over the job of the absentee. She found that she could do quite well on this new job. Meanwhile, a spare operative had been moved into the team to take over 15's job. When the absentee returned Operative 15 was not moved back immediately to her old job but was found a few odd batches of work by the chargehand to keep her occupied. She was upset about this and also about the waiting time she was accumulating while the chargehand found work for her. Her experiences are reflected in her pay packet. Another cause of earnings fluctuations for Operative 15 was her reliance on the investigator to keep her supplied with parts. He was an inexperienced worker and often failed to keep up with her. And when the investigator was waiting for work to be delivered from the cutting room delay for both was unavoidable.

Operative No. 14.

In Fig. 1 periods during which new styles were introduced into the units are shown by arrowed lines.[1] The influence of a change in style is clearly reflected in the earnings of Operative 14. Her earnings were well below her own average from the second to the sixth week of the period. With the introduction of a new style about the sixth week her earnings climb well above average. When the style changes again over the tenth and eleventh weeks earnings within a week or two slump below average again. Then there is a steady climb with the introduction of two new styles towards the end of the 19 week period.

The sudden slump in Week 15 is due to work starting on a new and difficult coat. This operative, in common with others,

[1] See note to Fig. 1.

was experiencing a period of uncertainty. Not only was a new style coming into production but the cloth was difficult to work and the operatives did not know the prices which were to be paid on the new style. There were rumours that all the work was to be retimed. Others said that average rates were to be paid because the run was to be short. There was general uncertainty about how work on the new coat was to be allocated, and about job prices. After one or two of the jobs had been retimed the prices were announced. Most of the operatives were then satisfied that they could "make their money." From then on wages rose for Operative 14, as for many others.

Operative No. 12.

The influence of changes in style in earnings is again illustrated in the earnings graph of Operative 12, although not so clearly as for Operatives 14 and 10. The earnings of Operative 12 also show marked variations on the same style. This operative, a machinist, undertook a great variety of operations. Her regular task, that of completing the sewing of the garment for final assembly, did not occupy her whole time, and she looked out for "bits and pieces" to keep her busy, either from the chargehand or from other workers who had too much work. "Bits and pieces" were not always available and in this event waiting time reduced earnings. And of course frequent changes of operation tended to slow down her work and break the rhythm, again with adverse effects on earnings. In weeks when there was a good run of her regular work, which depended upon the final assemblers' output, her earnings tended to be better. It might be argued that the operative could have planned her work better to iron out fluctuations, but this would have been very difficult since the supply of work to her depended upon the activities of other people, and they had similar problems. She was, like most of the other operatives, affected by the general uncertainty which surrounded the introduction of the new style in weeks 14 and 15.

Operative No. 10.

Like Operative 14, Operative 10's earnings seem to be affected most by changes in styles. These two operatives were amongst the more highly skilled workers in the unit and they

were rarely transferred to other work. But they performed operations which often changed radically when new styles were introduced. The slump in week 8 is explained by the effect of Operative 10's own absence at the beginning of the week. She remarked that she wished she had taken the whole week off since her work had been done for her by two other people, and she came back to find herself plagued by waiting time until the work began to flow again.

The precipitate drop in earnings in week 6 is accounted for by an error in assembly. Several coats on a new style had been completed before the error was noticed and all the work had to be undone and made again.

Operative No. 7.

The earnings of this operative were affected only slightly by most changes of style, but markedly by the style change of weeks 14 and 15. The other variations in her earnings are to be largely explained by the fact that she worked on an important sub-assembly operation. Her capacity to earn was affected by the efforts of about half the workers in the unit. Backs, fronts and sleeves were supplied to her by other workers, and she joined these together. Any hold-up in the supply of any of these would prevent her from going ahead. But drastic slumps due to occurrences like this were avoided because this operative, besides being a skilled machinist, was also a skilled assembler. And when work was held up on her machine she could usually, but not always, find some assembly work to offset the effect of waiting time on her earnings.

The difficulties of Operative 7 are perhaps best illustrated from our field notes. At the beginning of week 6 this operative told us that to "make her wages" (*i.e.* about £1.0.0. a day) she had to complete about sixty joining-up operations. She found that on the new style it was difficult for other workers to keep her supplied at this rate. She said that she could do her sixty easily enough if she got the parts, but if she did sixty this was too many for the assemblers who took the work from her to keep up with. To escape this dilemma she worked at the rate of sixty a day until supplies ran out, and then left her machine to help the assemblers clear away the work she had

accumulated for them. But, as she explained, her assembly work was slower than her machine work because of lack of practice, and she lost out. Week 6 is a good example of this.

It is impossible to trace clearly the effect of "sweating" on earnings. But the measures taken by the inspectors to prevent deterioration of quality, and the effect of these on earnings, can be shown. We refer again to Operative 15, particularly for weeks 5 and 6. This operative told us that she was pleased that a new style was coming into production as the unit had made this style before and she knew that she could "make her wages" on it. She started "sweating" to make up for the poor earnings of week 5, but soon found that she had a batch of work returned by the inspectors which she had to rectify. This kept her earnings low in week 6.

Operative No. 3.

This operative's earnings are fairly steady. In this she resembled her colleague Operative 2. These two young women worked together on a small part of the garment which changed little when styles changed. We have no account of either of them being transferred to other work when other operatives were absent. They did not rely, as did Operative 15 and others, on the output of other operatives. They collected parts direct, and after completing the operations on them the work was delivered straight to those on final assembly. We would say that such fluctuations as these two operatives experienced were due to waiting time. We often saw them waiting for parts to be delivered to the unit. It may also be the case that they used "dead horse" to even out their earnings, but we have no information about this.

Summary.

The system of incentive payments in Factory I was designed to reward each increment of output with an equal cash increment. To decide whether it did in fact do so would involve an attempt to disentangle *effort* from the whole complex of factors affecting earnings. This is impossible, especially in the absence of accurate measures. But it is clear that several factors combined to determine what earnings would be in any

one period of time. Unlike the factor *effort expended*, most of these factors lay outside the direct control of the workers. If these factors operated to depress earnings below what the workers considered fair and equitable, it would seem reasonable to expect that attempts would be made to modify their effects. In this workshop the modifying factors which we mentioned previously were not very effective in their operation. The system of allocating and "booking-in" work made "dead horse" difficult to operate. "Sweating" was largely defeated by the system of inspection. The informal and personal system of re-allocating work served to some extent to iron out fluctuations, but it also involved transfers, and these, as we have seen, sometimes had a depressing effect on earnings.

If it be accepted that in this workshop there were no gross fluctuations in *effort expended* sufficient to account for all fluctuations in earnings, it follows that the job prices did not accurately take into account the effect of other factors i.e. waiting time, changes in style, etc. Our material suggests that any attempt to do this would be doomed to failure from the start since so many of the factors affecting earnings are irregular in their incidence, and unpredictable, and this points to the need for "on the spot" bargaining to compensate for the effects of particular factors in terms of a general agreement about the price for the job.

It is possible to argue from the figures we have presented that the average earnings shown are adequate reward for work done and are the end product of a process of "taking the rough with the smooth." But this would not accord with a claim that time-study and job-evaluation take most of the guesswork out of job-pricing. It must also be remembered that the job prices are calculated from a datum (the industry minimum) which is itself the outcome of bargaining at a higher level and represents an evaluation of the worth of a contemporary semi-skilled worker, on time rates. All in all the material we have presented for Factory I would seem to argue a case that negotiated industry rates should take account of the factors we have discussed, or that provision should be made for "on the spot" bargaining. The first course is obviously

impracticable, the second involves devolution of control to the workshop, and provided this is desired by the workpeople themselves, it would seem to be a workable solution.[1] The account of Factory II which follows is of relevance to this general line of argument.

Factory II.

(1) *Product and method of production.*

In Factory II the investigator worked in a small workshop where engineering equipment was being assembled. Data on the earnings of eighteen semi-skilled workers were supplied by management. Three of the eighteen workers were *finishers*, whose task it was to fit assemblies into a container and finish the product. The other fifteen workers were *assemblers*, who built the assemblies. All the *assemblers* were young men in the age range 20–40. Two of the finishers were in this age range and the other was a man in his fifties. With one exception, all the workers had less than seven years of service with the firm and each of them had a varied occupational history.

A great variety of work passed through the shop. The product was usually made to customers' specifications and covered a wide range of sizes and designs, though there were some ranges of standard designs. The time taken to complete their work varied between one hour and thirty hours for the assemblers, and between one hour and ten hours for the finishers. But despite the wide variation in size and design, the main components for each product were similar, as was the sequence of operations. Each worker was expected to tackle any job allocated to him, and each was paid according to the same scale of wages and incentives. Broadly speaking, all the workers were at the same skill level, although the foreman allocated the work in the shop according to the special aptitudes of workers.

[1]Managers in the factory who kindly read and commented on this report, have emphasised that all that time-study can achieve is a reduction of the area of evaluation in job-pricing. Their main argument against "on the spot" bargaining as a remedy, is that a resort to bargaining as a means of settling day-to-day problems would be impracticable, and make rate-setting even more unreliable.

The assembly work, and the finishing work, was carried out with hand tools. The assemblies were built on conveyors, but this was only for convenience when moving them up the shop for finishing and testing; there was no mass-production sequence. Each assembly was built complete by one worker, or by two "working mates." Overhead cranes were used for lifting larger components in place, and for moving large assemblies up the shop. Thus there was no machine-pacing of the work. The output of the shop depended entirely upon the efforts of the workers, given an uninterrupted flow of component parts. The incentive bonus scheme was designed to maximise speed of assembly, consistent with a given level of quality being maintained.

(2) *The Incentive System.*

Incentive payments were based upon *allowed times*. The *allowed times* were arrived at after time-studies had been taken by ratefixers, but these *allowed times* were always the subject of negotiation on the shop floor. Often the "offer" made by the ratefixer would be amended after discussion involving the operative concerned, the shop steward, and the foreman. The *time saved on allowed times* was expressed as a *percentage bonus*. From this percentage a weekly cash increment was calculated. The formulae for calculating bonus percentages and cash increments will be described presently.

Table II shows the items which entered into the calculation of weekly earnings. Most of these items were the outcome of national negotiations covering the industry. Item *A* in Table II is the negotiated weekly time rate of wages, and it applies to the whole industry. The sum of items *B* and *C*, which is shown in column *D*, is the negotiated weekly piecework rate for the same class of workers. Items *A* and *C* were increased by 11/- as a result of the national award which took effect on 5.3.1956, at the time when the research was in progress, therefore item *D* was also increased by 11/-. But item *B* was not affected by the national award.

TABLE II

FACTORY II—RATES

	Daywork Rate	Piecework Rate		
	A	B	C	D (B+C)
PRE 5/3/56 ...	151/6	64/2	68/4	132/6
POST 5/3/56 ...	162/6	64/2	79/4	143/6

Note :

These rates apply to workers with 24 months service or more. Those with less service are regarded as trainees. All trainee rates are lower than those shown, but some allowances are made to compensate.

In calculating weekly earnings the following procedure was adopted. All hours spent on piecework in the course of the week were paid for at the piecework hourly rate $\frac{(B + C)}{44}$. All hours spent waiting for work, or on work declared to be *daywork*, were paid for at the hourly time rate $\frac{(A)}{44}$. The incentive bonus percentage was calculated according to the formula $127\frac{1}{2} \times \frac{\text{time allowed}}{\text{time taken}} - 100$.[1] Once the % bonus had been calculated for all jobs done in the week, the cash increment was calculated as follows : % bonus $\times \frac{\text{piecework hours}}{44} \times$ piecework datum (B).

Thus the three main items in the weekly wage calculation were :

i. Payment for daywork at time rates.

ii. Payment for piecework at piecework rates.

iii. Bonus for time saved on allowed times.

[1]The chief point about this calculation is that if the job is done in exactly the allowed time the rate of bonus earned is $27\frac{1}{2}$%. Translated into cash terms (i.e. as a proportion of the piecework datum (B)) this is the difference between the day-work rate (A) and the piecework rate $(B + C)$.

But there were other items in the wages calculation. Time spent waiting for the crane carried a bonus allowance of 45%, and a similar bonus was paid for delays due to factors beyond the control of the workers, e.g. faulty design, poor materials. In addition to these a shop productivity bonus was paid each week. This was expressed as a percentage increment to weekly earnings and was obtained by subtracting an agreed *average bonus* percentage from the *actual bonus* earned by the operatives each week (expressed as an average percentage for the whole department).

Table III is an example of a weekly wages calculation. This calculation is included because it contains most of the items enumerated above. It is typical save in one respect ; the daywork hours in the calculation are well in excess of the average for the shop.

TABLE III

TYPICAL WEEKLY EARNINGS CALCULATION OF A WORKER IN FACTORY II

Rate of bonus earned (time saved on allowed times) = 84%

Item	*Description*		*Payment*
1	Payment for daywork hours (20·1 hrs. @ 162/6)	=	74/3
2	Payment for piecework hours (27·9 @ 143/6)	=	91/–
3	Piecework bonus earnings $(84\% \times \frac{27{\cdot}9 \times 64/2}{44})$	=	33/11
4	Overtime allowance	=	7/5
5	Productivity bonus	=	9/10
	Gross earnings	=	216/5

Note : (*a*) This calculation is based on post 5/3/56 rates (see Table II).
(*b*) The daywork hours in this calculation are well above the average for the shop.

When considering the earnings records which were made available by the management it should be borne in mind that these do not necessarily represent the actual working experience of the workers each week. The earnings calculations are based upon times *booked-in* by the operatives, and operatives did not always book-in the actual times taken. If they considered, for example, that the allowed time on a job was "loose," i.e. that

it was easy to earn bonus on, they would often book-in less than the actual time saved, as a safeguard against possible rate-cutting.[1] They were thus enabled in many cases to build up a "bank" of hours which could be "cross-booked" onto "tight" jobs, or to adjust the proportion between daywork and piecework hours in the wages calculation.[2]

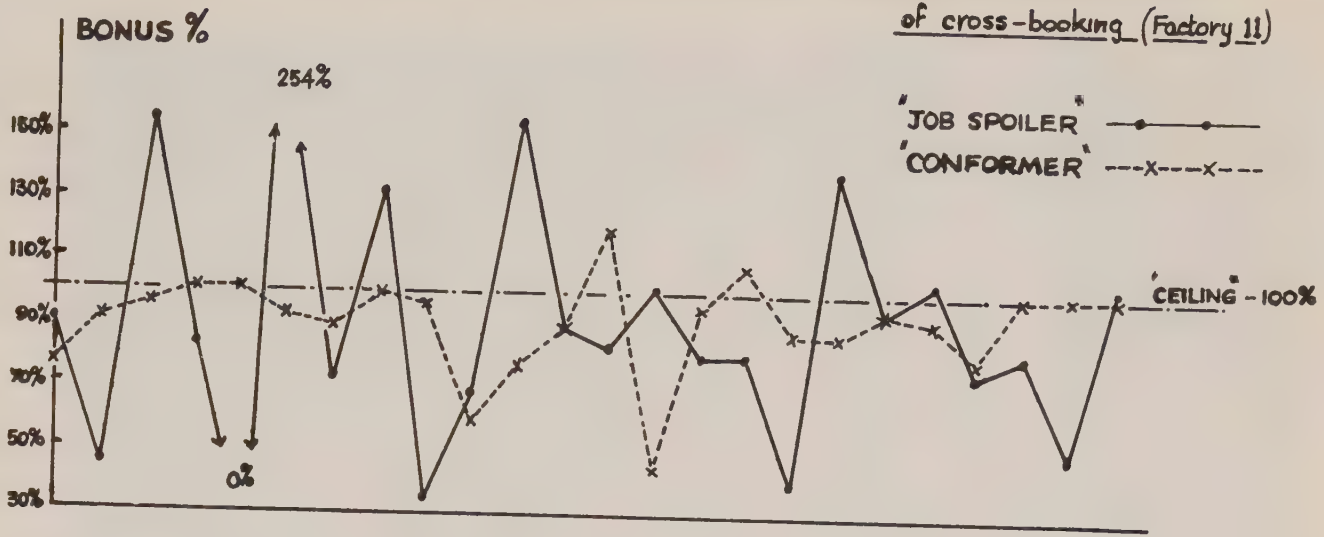
There were other ways in which the incentive scheme was manipulated so as to maximise earnings, and to even out fluctuations. We shall discuss these later. But it is already obvious that the earnings figures which we shall present below will not necessarily represent the true account of how time was spent in the shop. The fluctuations which these figures will show are probably much less violent than they would have been if they had been based on actual time spent rather than on actual times booked-in. Fig. 2 illustrates this point well. Here fluctuations in the earnings of a worker who conformed to the general pattern of behaviour in the shop are compared with the fluctuation in the earnings of a "job spoiler." The conformer normally kept under the bonus "ceiling," [3] as part of the defence against rate-cutting, and he "cross-booked." The job spoiler "booked straight." The difference in the range of fluctuations in the average hourly earnings of these two men is remarkable. It is likely therefore that had all the men in the shop "booked straight," the fluctuations in their average hourly earnings each week would have been more marked than in fact they were. Even so the attempt to iron out fluctuations by "cross-booking" fell short of complete success, and the earnings of all the operatives show fluctuations which are in some cases quite marked. To explain these fluctuations a wide range of factors has to be taken into account. We assume that the policy of "cross-booking" was pursued for two reasons; firstly, to safeguard against rate-cutting, and secondly, to

[1]There was an agreement between management and the union that a rate once set on any job could not be altered unless the method was changed. Despite this, fear of rate-cutting was often expressed.

[2]We have evidence that this practice is fairly common in engineering.

[3]With few exceptions, the workers considered that to book-in times saved which would earn bonus of more than 100% was to invite rate-cutting. We refer to this figure as the bonus "ceiling." If *actual* time saved was at a bonus rate of 100% plus, the surplus hours could often be held in reserve to be "cross-booked" to a "tight" job.

FIG.2 Comparison of "job spoiler" & "conformer" – illustrating stabilising effect of cross-booking (Factory 11)



stabilise earnings. It is not possible to say whether rates would have been cut if the policy of cross-booking had not been followed. We did not observe a situation where cross-booking was generally avoided. Only in isolated cases did operatives eschew cross-booking as a policy. As for the second reason, stability, it would seem that its relative failure must be explained by reference to factors which the operatives found difficult or impossible to control, and which therefore limited their scope for ironing out fluctuations.

(3) *Factors influencing the capacity to earn.*

(*a*) *Waiting time.*

Some account of the system by which the department was "fed" with components for assembly is necessary if the phenomenon of waiting time at Factory II is to be understood. The two major components for assembly came from sub-assembly departments. The small components for each order, wire, insulating materials, nuts and bolts, etc., were collected in "feeder" stores and these were drawn when the major components had been delivered to the workshop and the job had been allocated by the foreman. Work could not be started until *both* the main components were delivered, and the assembly could not be completed until a complete set of small stores was ready. The latter were supplied to the departmental "feeder" stores, either by outside firms, or by machine shops within the plant. Any hold-up in the supply of components either from "feeder" departments or to the "feeder" stores might either delay the start of a job or cause a hold-up at some stage in the course of assembly and finishing. If operatives were kept waiting for work this was usually due to one of two things ; to shortage of orders, when no work at all was delivered from the "feeder" departments, or to lack of co-ordination of flow of main components from "feeder" departments or of small parts into the stores. The remedy lay outside the control of the operatives in the shop. The *finishers* were supplied with containers by the container department and with assemblies by the assemblers. And they also "drew" sets of small stores from the departmental stores. Like the assemblers they could not begin work until *both* main components were delivered and

they could not finish until they had a complete set of small stores. It is notoriously difficult to co-ordinate work flow on batch production of this kind, especially when the assembly is done mainly by hand. Inevitably, there are interruptions in the flow and it is at the final assembly points that the effects of these interruptions are most severely felt. So, if workers were short of components there was little they could do except wait, unless they could keep several jobs incomplete and then complete them over a short period when parts became available. But this was seldom done. To do it would have meant overloading the department, and this might have had the effect of aggravating the dislocations in the work flow over the whole factory. In any case the operatives disliked having incomplete work around. They were not permitted to book time to incomplete jobs and if there were many of these bonus payment was thereby postponed.

(*b*) *Waiting for cranes.*

Since a number of sub-departments were served by the same overhead crane, operatives were sometimes held up "waiting for a lift." Time booked in this way was paid for at 45%. But here again the operative on a "loose" job might not book-in time spent "waiting for crane." Instead, it was sometimes possible to book-in at the rate of, say, 90% for the whole job, as if no delay had been experienced. Workers were conscious that they were doing this but we heard no one mention that this also served to keep up the shop bonus average, and in this way increased the productivity bonus. The tendency to keep the shop bonus average low by operating a bonus "ceiling" was also offset by this practice.

(*c*) *Design and faulty drawings.*

When delays were occasioned by drawing office errors it was sometimes possible for the operative to negotiate with the ratefixer an "on the spot" allowance which offset the effects on his earnings of these delays. But a run of such difficulties might result in lowered earnings.

(*d*) *Faulty materials.*

Sometimes the components delivered to the worker were faulty and had to be taken back to the "feeder" departments for rectification. Delays of this kind were compensated for as follows. All time booked-in was counted as piecework time and a bonus of 45% was paid. But here the operative had a degree of control since if the time were "loose" he could ignore the delays, and book in a time taken which would pay 90% (say). This would mean a reduction of potential "time up the cuff" [1] thus limiting opportunities in the future for offsetting the effects of waiting time and "tight" times.

(*e*) *Allocation of work.*

The allocation of work within the shop was in the hands of the foreman. He favoured certain of those whom he considered his best workers. Complaints were often heard that work was not evenly distributed amongst the operatives, nor was it distributed evenly over time to the whole shop. Since there were "tight" or "loose" rates the effect was to create a discrepancy between the opportunities to earn bonus as between the men, and to affect the capacity of the whole shop to stabilise its earnings. Of course, pressures were often brought to bear on the foreman to reform his method of allocating work, so as to give everyone the same proportion of "good" or "bad" work, and sometimes he would yield. To this limited extent the operatives controlled the allocation of work within the shop. But neither they nor the foreman controlled the flow of components into the shop.

Another factor affecting the stability of weekly earnings was the rule that no bonus would be paid on a job until it was complete. If, for example, a man had worked for most of a week on an assembly but had not completed it by Friday, he would draw no bonus for the job that week. He would not lose by this, for next week his bonus average would rise. But it would mean a slump in his earnings for the week in which the job was left incomplete. To offset this effect it was sometimes possible to book the work as complete and finish it on Monday, but, for

[1] This was one of the many expressions used to describe time which had been saved on loosely rated jobs.

reasons already mentioned, this was difficult to arrange and uncertain in its effects.

(4) *Modifying factors.*

Under this heading we sum up the techniques which were employed by the operatives to offset the effects of the factors which influenced their capacity to earn. Most of these have already been briefly touched upon. There was the technique of "cross-booking." The responsibility for declaring the time taken to complete a task rested with the operative concerned. When the allowed time was "loose" most operatives were careful not to declare excessive time saved as this would expose the "looseness" of the rate. Most operatives recognised a bonus "ceiling" of about 100%. Time not booked-in on loosely rated jobs could usually be booked over on to a tight job, or kept "up the cuff" until the occasion arose to use it either for this purpose or to offset the effects of waiting time. For example, an operative who had been kept waiting for the crane would not claim the 45% allowance but would use "banked" time to make the job pay, at a higher rate of bonus.

Since the existence of loose rates was essential to the success of the "cross-booking" technique, the operatives tried always to strike a hard bargain with the ratefixer when jobs were being timed, and to employ all sorts of devices to spin out the work during actual time-studies.

Thus the allocation of time and the payment for it was to some extent within the control of the operatives. And this meant that they could sometimes afford to take things easy when they were on loosely rated jobs without serious financial consequences.

(5) *Average hourly earnings.*

Table IV shows the range of fluctuation of average hourly earnings for 15 assemblers and three finishers over a twenty-five week period. The average hourly earnings in a week have been computed by dividing weekly gross earnings by hours worked. In calculating the overall average for each operative we have ignored the effects of the wage increase which took effect in week 10. This increase did not affect the bonus datum and seems to have made little difference to the earnings of the

operatives in this shop. This emerges more clearly from Fig. 3 which shows graphs of average hourly earnings for four of the operatives—three assemblers and one finisher. This figure also shows bonus percentages and daywork hours for these operatives.

TABLE IV

FLUCTUATIONS IN AVERAGE HOURLY EARNINGS

Operative	Average hourly rate		Average	Range	
	Lowest	Highest			
1	4/1	5/5½	4/6	1/4½	Assemblers
2	4/1	5/4	4/5½	1/3	
3	3/3	4/6½	3/11¾	1/3½	
4	3/11¾	5/5¾	4/5½	1/6	
5	3/11½	5/4¾	4/5	1/5½	
6	4/–	4/8	4/4¾	8d. *	
7	4/0½	4/11½	4/6	11d.	
8	3/3	4/7¾	3/10½	1/4¾ †	
9	3/6	4/9	4/3½	1/3	
10	3/3½	3/9½	3/5¾	5¾d. *	
11	3/2	7/8 ‡	4/1½	4/6 †	
12	3/4	5/10	3/11½	2/6 †	
13	3/6½	5/0¾	4/3¾	1/6½	
14	3/2	6/10½	4/6	3/8½	
15	3/9	5/6½	4/6½	1/9½	
16	3/11¾	4/10½	4/3¾	10¾d.	Finishers
17	3/11¾	4/10½	4/3¾	10¾d.	
18	4/–	5/0½	4/4½	1/0½	

* For 10 weeks only.

† On trainee rates for part of period—which widens range and decreases average.

‡ A "freak" figure, due to booking incomplete work.

Note: Operatives who habitually "worked mates" are bracketed together.

We now examine the graphs in Fig. 3 in more detail.

FIG. 3 FLUCTUATIONS IN EARNINGS ETC. Factory 11

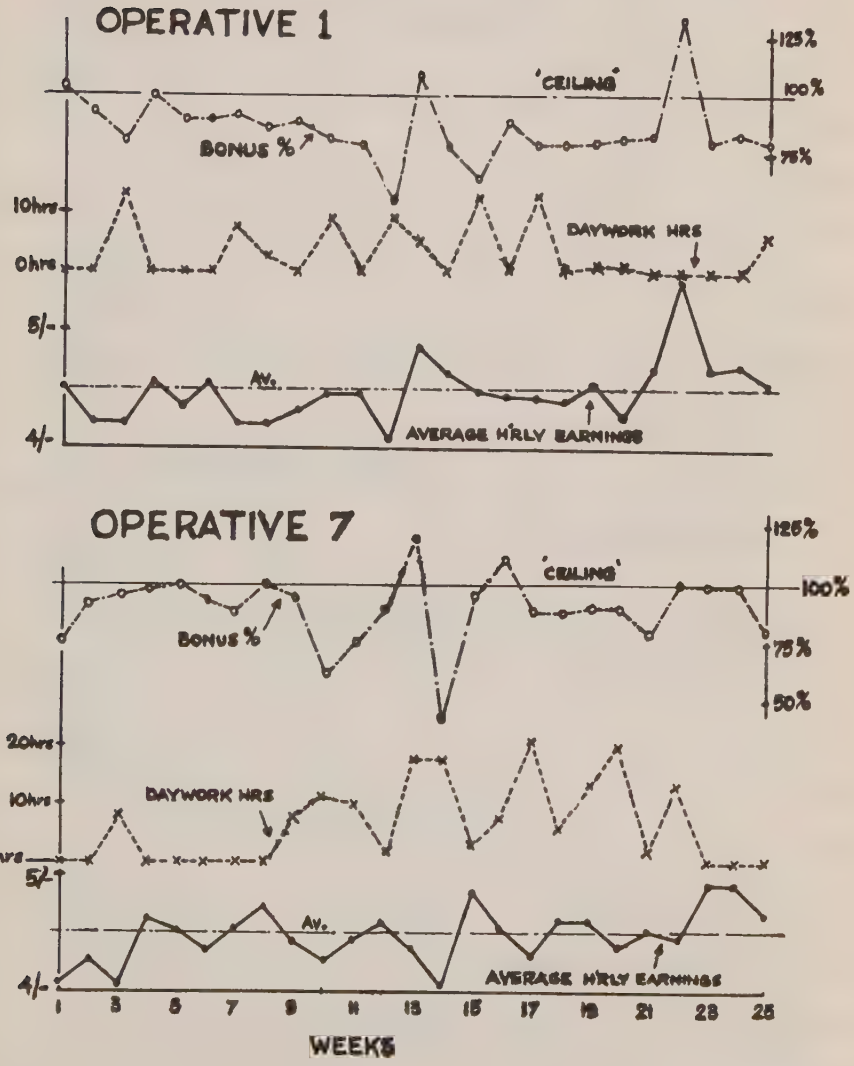
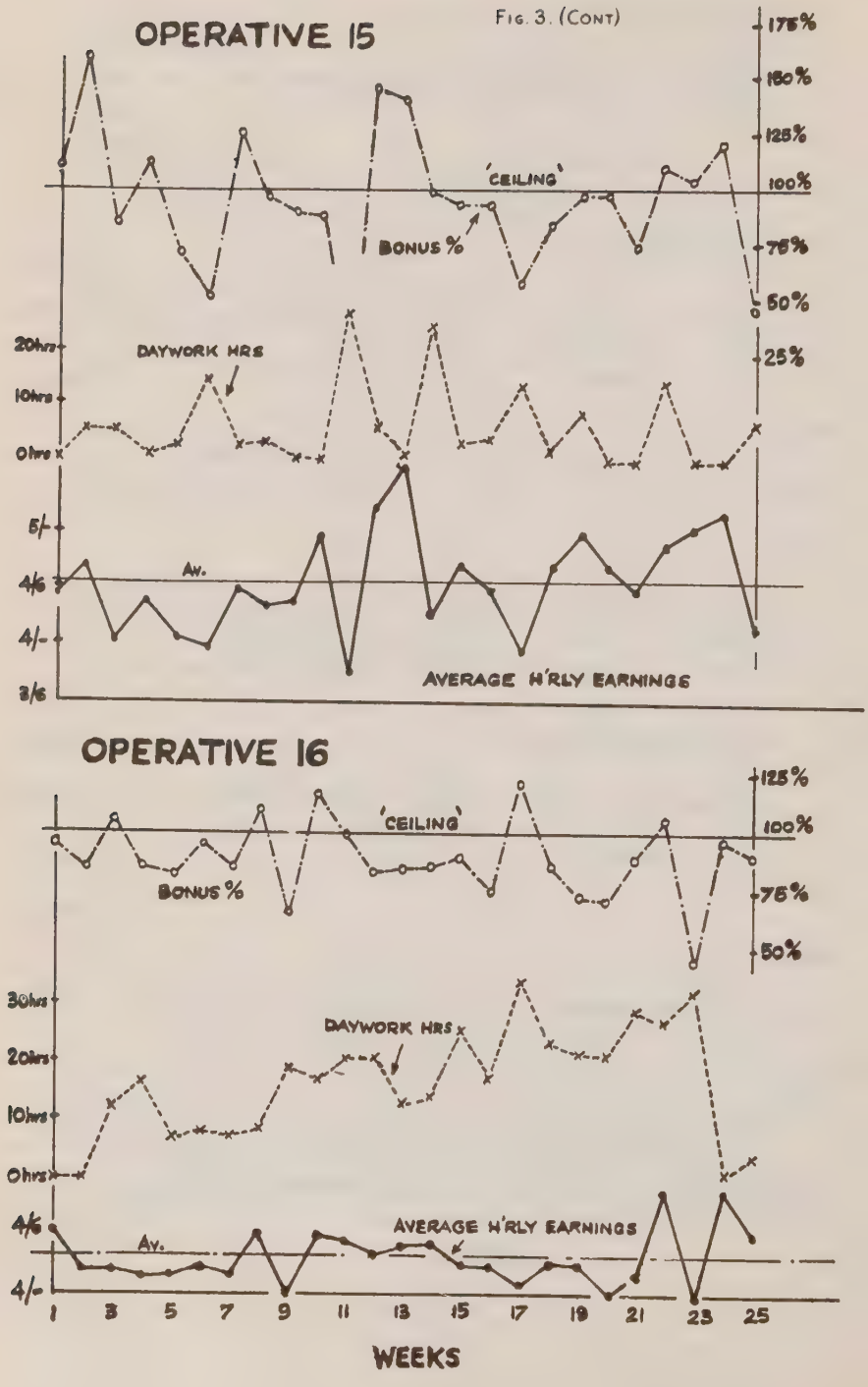


FIG. 3. (CONT)


OPERATIVE 15
175%
150%
125%
100%
75%
50%
'CEILING'
BONUS %
20hrs
10hrs
0hrs
DAYWORK HRS
25%
5/-
4/6
4/-
3/6
AV.
AVERAGE H'RLY EARNINGS
OPERATIVE 16
125%
100%
75%
50%
'CEILING'
BONUS %
30hrs
20hrs
10hrs
0hrs
DAYWORK HRS
4/6
4/-
AV.
AVERAGE H'RLY EARNINGS
1
3
5
7
9
11
13
15
17
19
21
23
25
WEEKS

Operative 16 (*finisher*).

With the exception of the two "boom" weeks 22 and 24 and the two "slump" weeks 9 and 23, this operative's earnings are fairly steady over the period. This is remarkable when the large amount of daywork is considered. Daywork reduces the opportunity to earn piecework bonus and, given "straight" booking, differences in daywork from week to week would be exactly reflected in average hourly earnings so long as the bonus percentage remained steady. That this operative has ironed out his earnings' fluctuations while keeping his bonus below the ceiling for most weeks indicates his great skill in the technique of "cross-booking."

Week 17 is a good illustration of the way the technique was used. There was a sharp rise in daywork hours. This was compensated for by booking in high bonus, well above the "ceiling," and a drastic slump in earnings was avoided. It was possible to book-in high bonus if rates were "loose," if rates were "reasonable" and tremendous effort was exerted, or if there was "time up the cuff" saved for an eventuality such as this. We have no record that this operative worked any harder in this week than in others, and we conclude that time from "up the cuff" was used. In weeks 9 and 23 the operative was unable to avoid a slump in earnings occasioned by sharp rises in the amount of daywork. It is possible that in these weeks there was a run of tight work or an absence of "time up the cuff," or both together. We have no evidence that this was in fact the case. We were unable to follow the "booking-in" practice of all the operatives from day to day. There were two reasons for this. Firstly, we had a job to do in the shop which limited research time and, secondly, operatives tended to conceal the details of their "booking-in" practices. Our information about "booking-in" practices comes from conversations overheard and participated in, and responses from selected informants.

From the general custom of the shop on these matters we infer effects on earnings, where detailed information is lacking. Close examination of the graph shows how effective "cross-booking" is as a stabiliser. We shall return later to consider the question of whether it "pays" in terms of high average earnings and leisure time won. Now we must consider the earnings of an operative who did not conform to shop "booking-in" conventions.

Operative 15.

This operative was considered to be a "job spoiler." He cared little about the consequences of exposing "loose" times by booking-in high bonus percentages. An examination of the bonus, daywork hours, and average earnings graphs of this operative shows what happens when the bonus "ceiling" is disregarded and "cross-booking" is little used. Take week 11 for example. In this week there is a marked slump in bonus earnings combined with a sharp rise in the amount of waiting time. This has a catastrophic effect on earnings. In the absence of detailed evidence, we presume that in this week the operative had a run of tight jobs during the time he spent on piecework, or that he had incomplete work at the end of the week. It also seems likely that this operative faced a similar situation in week 17, another week of low bonus and high daywork. But when daywork is low and bonus is high, earnings soar, as in weeks 12 and 13. When the graphs of operatives 15 and 16 are compared it looks as if "cross-booking" achieves a balance of advantages to the operative. Despite the large amount of daywork experienced by operative 16 his average earnings are not much less than those of operative 15. And operative 15 certainly worked much faster than operative 16, being regarded in the shop as a "tear-arse." And the earnings of operative 16 are much more stable over time. Operative 15's earnings are all over the place. All this is apart from the imputed gain from normal observance of the "ceiling," which is to safeguard against rate-cutting and preserve the "loose" rates which make cross-booking possible. These observations are given further point if we examine the figures of another "conforming" operative.

Operative 7.

This operative, an assembler, rarely exceeded the bonus ceiling. He did so only twice in the 25 week period, week 9 and week 16. On both occasions high bonus compensated for high day-work hours. Week 10 shows a drop in average hourly earnings associated with low bonus percentage and high daywork indicating perhaps a run of "tight" jobs associated with a shortage of hours "up the cuff." This operative had less day-

work than operative 1 so that the problem of ensuring stable earnings was less difficult.

Operative 1.

There are two features of operative 1's graphs which need special mention.

1. His bonus percentages are low in comparison with the other three.
2. He has much less waiting time.

These features are explained by the fact that this operative often worked on special work demanding high quality. This work was not paid according to the incentive scheme. A standard percentage bonus was paid irrespective of the time taken. This percentage was usually much lower than the ceiling—say 80 or 85%. No "cross-booking" was possible, of course, in a week wholly occupied with special jobs, but if the operative combined work on "specials" with ordinary piecework jobs, it was possible to make tight jobs pay by declaring time to have been saved which was actually saved on the special jobs. The fact that the operative worked on "specials" also explains his low day-work hours.

This operative, as he often told us, did not think it wise consistently to book above the "ceiling." He added that in his career he had experienced rate-cutting, and we would say that he kept his earnings steadier than they would otherwise have been by some "cross-booking." The high bonus earnings in week 22 which, together with absence of daywork, sent his earnings soaring, were probably due to incomplete work from the previous week being "booked-in."

In Factory II the operatives, with one or two exceptions, operated the "cross-booking" technique to stabilise earnings and to safeguard against rate-cutting. In doing so they compensated for the uncertainties in work flow typical of batch production shops. That they were able to exercise the control they did was probably due to the existence of strong Trade Union organisation in the workshops, and a tradition of bargaining about rates at workshop level.

Summary.

In Factory II, a genuine attempt was made by "on the spot" bargaining to take account of all the factors likely to influence the level of earnings, and thus to compensate the operatives for losses due to factors beyond their control. But in this case the incidence of waiting time which, by agreement, was paid for at time rates, introduced an unpredictable factor. To offset the effect of this factor the "booking-in" procedure was manipulated by the workers. Successful manipulation demanded the existence of "loose" rates and these were preserved by the expedients of the "bonus ceiling," cross-booking, and hard bargaining with the ratefixers.

Factory III.

(1) *The industry.*

This section of the clothing industry is subject to periods of boom and slump, of both seasonal and longer duration. Most of the firms are small, and competition between them is intense, particularly in periods of slack trade. The industry was experiencing such a slack period when the study was made. Traditionally labour within the industry is mobile. Workers have long been accustomed to move from firm to firm in search of work or of higher wages.

(2) *The method of production.*

Between 50 and 100 people, including both men and women, were employed at Factory III, on the assembly of garments. Two distinct processes were involved in making most of the garments. There were two groups of skilled workers in the factory, referred to below as *assemblers* and *machinists*. An assembler worked in conjunction with a machinist making each garment all the way through. The garment passed back and forth between assembler and machinist about three times in the course of making up. A machinist worked in conjunction with from one to five assemblers. Each assembler would always give his work to the same machinist. The majority of garments were made under this system, referred to below as *Production System A*. A small proportion of garments was designed in such a way that they could be made up entirely by machinists.

This system is referred to as *Production System B*. The majority of machinists worked mainly on Production System A but made a few garments each week under Production System B. A minority worked solely under Production System B. Production was carried out under the supervision of one manager. The manager was in charge of allocation of work, and was assisted by a chargehand. There was no organized system of work flow. Garments to be assembled were delivered from the cutting room in bundles of from eight to twenty garments. Work made under Production System A was given to assemblers who went to the manager or chargehand when they needed it. Work made under Production System B was given to machinists by the chargehand.

(3) *The incentive scheme.*

The incentive plan is simply described as a straight piecework scheme based on a Price List. No discrimination was made between men and women. The Price List was first drawn up some forty years ago by a committee comprising lay Union Executive members and Employers' representatives. The list has been revised throughout on two or three occasions since, but without major alteration. The list was devised by breaking down the operations going into the making of a garment into small component parts. Over the years styles have multiplied and this has led to a further breakdown of operations and the addition of new items. The list now comprises about 900 items, about 350 applying to machinists and about 550 to assemblers.

The last revisions to the price list included an addition of a percentage bonus on all prices and of a percentage cost-of-living bonus to be added to prices. The latter does not operate automatically with changes in the cost-of-living index, but is the subject of negotiation. An increase of $7\frac{1}{2}\%$ was granted while the study was in progress and this came into effect during the last four weeks of the study (weeks 23 to 26 in the Tables and Graphs).

Characteristically, the clothing industry is subject to frequent changes in style and design. New materials are also being frequently introduced. Garments made at Factory III were usually in a fairly heavy weight or a fairly light weight

cloth. The difference between heavy and light weight cloth is catered for by a difference in bonus, five per cent. extra being paid for all work on heavy cloth. Materials made up at this factory ranged from beaverteens, twills, and cashmeres, to crepes and nylons in a great variety of colours. Allowances per garment for various types and colours have been negotiated by the Trade Union and the Employers. Similar allowances exist for differences in size of garments.

Operatives at Factory III claimed that in many ways the price list operated to produce unequal cash reward for equal effort. There were, they said, three anomalies of this kind.

(*i*) Prices of items were such that machinists could earn more than assemblers. There is no evidence that they worked harder. If anything the reverse was true.

(*ii*) Light work paid better than heavy work. Also allowances for colours and cloths were seldom commensurate with the extra effort involved in making them up.

(*iii*) Some designs of garment paid better than others, because some items in the list paid better than others. This was partly the result of the attempt to compare 900 different items valued as low as ¼d. or ⅛d. It was partly the result of the traditional policy of paying for length of sewing, regardless of the amount of handling involved in the sewing of an item. One outcome of this policy was that garments designed to be made under Production System B paid better than those under Production System A. The main reason for this is that under Production System B the garment did not have to pass back and forth between assemblers and machinists.

The anomalies in the price list meant that unequal cash reward for equal effort was built into the incentive system. According to the operatives, the anomalies were so great that extra effort could seldom offset the effect of "bad" work in any one week. The anomalies thus caused earnings to fluctuate. But to see the actual relation of the incentive system to fluctuations in earnings, it must be seen operating in the factory where mistakes, accidents and sickness occur and where market factors frequently cause an uneven work flow.

(4) *Factors affecting the capacity to earn.*

From the operative's point of view, given steady effort, there were four main factors affecting his capacity to earn.

(*a*) *Materials worked on.*

Materials varied in style and colour, making them more or less difficult to work on. Owing to the price list some of these materials were easy to make money on, while others were more difficult. From time to time, the cloth which the operatives had to make up was difficult to work on because of faults in it. There was no formal provision in the price list for extra payment in the case of faulty materials, and only very rarely was payment made.

(*b*) *Design or Style of Garment.*

Different designs of garment were more or less easy to make money on. Again the most important reasons for this derived from the price list as explained above. Though the "good" items in a garment often offset the "bad" ones, this was by no means always the case. Again garments made under Production System B tended to pay better than those made under Production System A. Quite frequently there were faults in the cutting of a garment. The operatives then had to spend time fitting and trimming the faulty cuttings or in getting recuts. They were not paid for this work, and so faults in cutting tended to depress their earnings.

(*c*) *Length of Run.*

If there was a good long run on one garment, operatives could work up speed on it. If they were constantly changing from one style to another, or from one material to another, this slowed down their pace of work. At Factory III a tremendous variety of styles was made and runs were usually short. Operatives usually made several styles each week. During time of depressed trade, when firms were competing desperately for orders, many sample garments and small orders were made, and runs were very short. A slightly different factor was the size of the bundle of garments which the operatives were given. Operatives generally worked on all the garments in a bundle at once, doing each operation on each garment in turn. A small

bundle made in this way took almost as long as a bigger bundle. In addition to length of runs becoming shorter in trade depression, the size of the bundle became smaller.

(*d*) *Waiting time.*

There was no payment in this factory for time spent waiting for work, unless the total weekly wage fell below the guaranteed union minimum wage. Many operatives worked part-time and in their case the minimum wage agreement did not apply. The wages of full-time operatives tended to be much higher than the minimum, hence much waiting time could occur without any payment being made. Waiting time was caused in periods of slack trade by a contraction in the total amount of work made.

Even when they are short of work, firms in this industry try to retain their labour force. They are then able to deliver quickly the orders they have in hand, an advantage since competition with respect of delivery dates becomes fiercer in periods of trade depression. Furthermore, they wish to avoid losing labour to competitors and to be in a good position when trade begins to get busy again. The effect of this policy on operatives was that they spent time waiting for work, since there was not enough to keep them all fully occupied. Another cause of waiting time in this factory was the breakdowns in the small assembler-machinist production units. For example, during the fieldwork, one machinist was absent. Another machinist then had to take on an extra assembler because of the absence of the machinist who regularly sewed her work. The machinist could not cope with the extra work, and kept her assemblers waiting. One assembler was kept waiting more frequently than the others, with a consequent drop of her earnings. Several quarrels occurred between this assembler and the machinist, and the former finally left the factory. An operative who was kept waiting for work in this way would open a fresh bundle of work if he had one. If not, he would go and ask for more work. Sometimes it was given to him, but other times, because work was so scarce, it was refused. Under Production System A, operatives usually worked on two bundles of work concurrently, one being with the assembler

and the other with the machinist. The earnings of an operative who had to work on three instead of the accustomed two bundles of work usually dropped.

These were the four main factors affecting an operative's capacity to earn. Their occurrence was beyond the direct control of operatives, and the effect on earnings was to make them fluctuate. But there are three other factors under the formal control of management, which the operative tried to control informally, in order to even out the fluctuations, or to raise his wages. We have called this type of factor a "modifying" factor.

(5) *Modifying Factors.*

(*a*) *Quality of work.*

The quality of work of different operatives varied. By producing work of low quality, an operative could get through a larger amount and so increase his earnings. Poor quality work sometimes passed the inspectors. In busy times, when there was a rush of orders, poor quality work passed relatively easily. But even in periods of slack trade, poor quality work did go through. In Factory III, some operatives were known to produce work of poorer quality than others at all times. These people were sometimes called "killers." "Killing" or "murdering" work resulted in some cases in considerably higher wages. There was a greater tendency to kill "bad" work than "good" work. In some circumstances, management connived at "killing," because work "murdered" was work made quickly.

(*b*) *The "dead horse" system of wage payment.*

During the week, work was given out to operatives, and was noted down in a book kept by the chargehand. At the end of each week, operatives were paid for a certain quantity of work, including work which had already been finished, and some which had not yet been finished. This latter was known as the "dead horse." It varied from week to week, not only because of variation in the quantity of work given out, paid for, and not yet finished, in any one week, but also because the operative might still not have finished work

which he had been given in the previous week. This latter factor meant that "dead horse" might accumulate over the weeks. Management's stated reason for the use of the "dead horse" system was that it evened out wage fluctuations : by providing wages for operatives who had been absent for one or two days in a week, thus spreading the loss in earnings over several weeks ; and also by rising in weeks when work was "bad" and falling when work was "good." There is no doubt that "dead horse" did fulfil these functions to some extent, in the first instance increasing, and in the second decreasing, fluctuations in average hourly earnings. But in doing so, it placed the operative in a debt relation to his employer. It is interesting to note that this is the exact opposite of the financial relation in Factory II where it was the worker who had "time up the cuff," that is, work he had completed but had not yet been paid for. "Dead horse" had other effects at Factory III. Since operatives felt under an obligation to pay off their debts, "dead horse" was a means of retaining labour. Though it may have reduced fluctuations over the period of a week, over longer periods it resulted in greater fluctuations. While trade was busy, an operative would try to get a lot of work on "dead horse," as this would raise his wages. Then, when trade began to get slack he would work off this "dead horse." In this way he artificially raised his wages in times of busy trade, and lowered them in times of slack trade. In this factory, "dead horse" had different effects for different people. Operatives who had high earnings did not have a lot of "dead horse" booked in busy periods. "Dead horse" thus redistributed work from the higher earners to the lower earners. It did so in another way. Lower paid operatives worked with, on the average, between one and two weeks "dead horse." In very slack periods, when there was little work coming through the factory, the people with "dead horse" were given enough work at the beginning of the week to cover their wages for the week. Paradoxically, they were given it all at once because they would not be able to make it that week, as they were working on their "dead horse." But still, they had to have it if they were to have any wages at all at the end of the week. The higher earners had very little "dead horse," so they were not

given their work all at once at the beginning of the week, but only as they finished their old work. However, when they were ready for work, all the work had sometimes been given out. They had to wait for work, while bundles were lying unmade on the tables of the low earners. In this way again "dead horse" redistributed work between the low and the high earners. The redistribution of work enabled the low earners to make enough wages to stay at the factory, and so enabled the employer to retain his labour force. The operatives had some control over "dead horse," since the amount of the weekly wage of each operative was frequently fixed in a private interview between the operative and the manager. The operative could complain if he thought that his wages were not high enough, and might succeed in persuading the manager to let him have more "dead horse."

(*c*) *Allocation of work.*

A great variety of work in different materials and styles was made at Factory III. As has been explained, these were more or less easy to earn money on. The way in which the work was allocated to different operatives thus affected their earnings. In a similar way, the allocation of different sized bundles affected earnings. Waiting time did not fall equally on all people, as we have explained in the section on "dead horse." Again, some of the operatives were more likely to "kill" work than others. All these can be considered as allocational factors affecting earnings. We do not want to go into the pattern of allocation here as this would involve a detailed sociological analysis. We will therefore just make two points. The manager had to allocate work primarily with market considerations in mind. Getting out orders quickly and keeping his labour force were more likely to determine his actions than trying to allocate work in a way which operatives thought was fair. When these two principles conflicted the former was much more likely to be followed. Secondly, there was in the factory no formal system of work allocation. All work was allocated personally either by the manager or by the chargehand under the direction of the manager. Under these conditions, when allocation of "good" and "bad" work caused

differences in earnings between different operatives, many accusations were made of favoured treatment of some operatives by the manager.

(6) *Comparison of average hourly earnings of operatives at Factory III.*

It is not possible to relate each of the fluctuations in any operative's earnings to any one of the seven factors listed, as several of them were usually operating simultaneously. Most of the graphs reflect the slackness in trade during the period of the study. Work was short, particularly in the first ten weeks, and this resulted in waiting time. Machinists suffered more from the shortage of work than assemblers, because the shortage of work was greatest for the garments made under Production System B. There were frequent changes in design of garment. Prices tended towards "bare union" rates. Light work was very scarce and for a period from around week nine to twenty nearly all operatives, including those who specialized on light work, were doing heavy work.

(7) *Average hourly earnings of assemblers compared.*

Average hourly earnings of assemblers for the 26-week period from January to July, 1956, are shown below in Table V, together with the range between the weeks when average hourly earnings were highest and those when they were lowest. Average hourly earnings are used as these make it possible to compare the earnings of part-time and full-time operatives. They are calculated in all cases by dividing the total earnings by the total number of hours worked in the relevant period. Freak weeks, with particularly high or low average hourly earnings due to the transfer of wages from one week to another through "dead horse" after absence, have been excluded from the range (but not of course from the average for the 26 weeks). Table V shows a considerable variation in the average hourly earnings and in the range. The highest earner earned almost four times as much as the lowest. The highest range was nearly eight times that of the lowest. Much of the differences in average hourly earnings can be attributed to differences in effort and skill. But work allocation and incidence of waiting

TABLE V

AVERAGE HOURLY EARNINGS AND RANGE OF ASSEMBLERS EMPLOYED DURING THE PERIOD JANUARY TO JULY, 1956

Number of Assembler	Lowest average hourly earnings	Highest average hourly earnings	Average for 26 weeks	Range	No. of weeks worked
1	5/3½	8/8	7/7½	3/4½	18
2	4/6¾	6/1	5/2¼	1/6½	15
3	3/9¼	5/5½	4/9	1/8¼	26
4	4/3	5/2¾	4/7	11¾d.	16
5	3/7¾	5/9	4/6½	2/1¼	26
6	3/7¼	5/9¼	4/4¾	2/2	26
7	3/7	5/–	4/3½	1/5	17
8	3/5½	4/11	4/3	1/5¾	26
9	3/–	5/7	4/2¾	2/7	26
10	2/11¼	4/10¾	4/0¼	1/11½	26
11	3/1¾	4/7½	4/–	1/5¾	26
12	3/1¾	4/6¾	3/10	1/5	26
13	2/3½	5/3	3/9¾	2/11½	23
14	3/6½	3/11¾	3/8½	5½d.	26
15	3/3	4/7½	3/8½	1/4½	14
16	2/3¾	4/8½	3/7½	2/4¾	26
17	2/2½	4/11¼	3/7¼	2/8¾	26
18	2/9½	4/1½	3/6½	1/4	16
19	2/5½	4/7	3/5½	2/1½	26
20	2/2	4/7½	3/5¼	2/5½	24
21	2/8	3/2½	3/4¾	6½d.	9
22	2/0½	4/0¼	3/4¼	1/11¾	16
23	1/3¾	4/3	3/4¼	2/11¼	18
24	2/0¾	4/3	3/3½	2/2½	18
25	2/2¾	3/9¾	3/1¾	1/7	24
26	2/4½	4/2¼	3/1	1/10	24
27	1/11	4/1¾	3/0½	2/2¾	26
28	1/11½	3/9½	2/11½	1/10	26
29	2/0¾	3/4½	2/10	1/3¾	26
30	1/10¾	3/5½	2/10	1/6¾	26
31	2/0½	3/7½	2/9¾	2/7	26
32	1/10¾	3/3	2/9	1/4½	25
33	2/0½	3/1½	2/8	1/1	26
34	1/11¼	3/8¾	2/7½	1/9½	16
35	4¼d.	3/1	2/1½	2/8¾	24

Notes :

Only operatives employed for more than 6 weeks are included in this table.

When calculating the range, weeks with exceptionally high or low average hourly earnings have been omitted when this is due to transfer of earnings from one week to another after absence, under the "dead horse" system.

time played an important part. The variation in the range is not so easily explicable. Even if the range is considered in proportion to average hourly earnings, there is still considerable variation. Some of the high ranges can be partly explained by difficulties in making heavy work experienced by operatives who usually work on light work. Allocational factors, however, account for a considerable part of the difference.

A similar argument can be applied to the earnings of machinists though no figures are given here. The acute shortage of work under Production System B, as well as the change from light to heavy work, affected different machinists differently. Again, however, allocational factors are important.

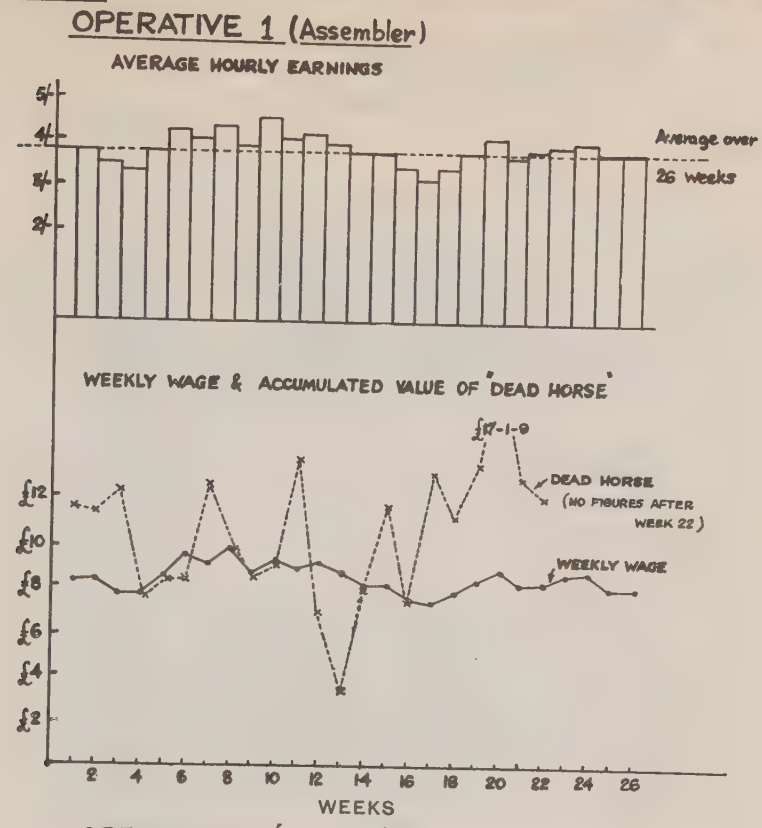
(8) *An analysis of weekly fluctuations in average hourly earnings for five operatives.*

Graphs are presented below (figs. 4 and 5) for five operatives showing their average hourly earnings for twenty-six weeks during the first six months of 1956.[1] Gross weekly earnings are given for three of these operatives, and the accumulated value of "dead horse" at the end of each week for two. The twenty-six weeks included four statutory paid holidays. In week 10 there were two days holiday, and in weeks 17 and 18, one. Holiday pay and holiday hours have been omitted from all calculations of average hourly earnings. For Operatives 1 to 4, holiday pay has been included in the weekly wage. For Operative 5, it has been omitted.

The graphs show that there was considerable fluctuation in average hourly earnings from week to week, even with the ironing-out effect of "dead horse." The graphs are taken in turn and an attempt made to relate the fluctuations in earnings over time to some of the seven factors listed above.

[1]The numbering of these operatives does not correspond with the numbering in Table V.

FIG. 4


OPERATIVE 1 (Assembler)
AVERAGE HOURLY EARNINGS
5/-
4/-
3/-
2/-
Average over
26 weeks
WEEKLY WAGE & ACCUMULATED VALUE OF "DEAD HORSE"
£17-1-9
£12
£10
£8
£6
£4
£2
DEAD HORSE
(NO FIGURES AFTER WEEK 22)
WEEKLY WAGE
2 4 6 8 10 12 14 16 18 20 22 24 26
WEEKS

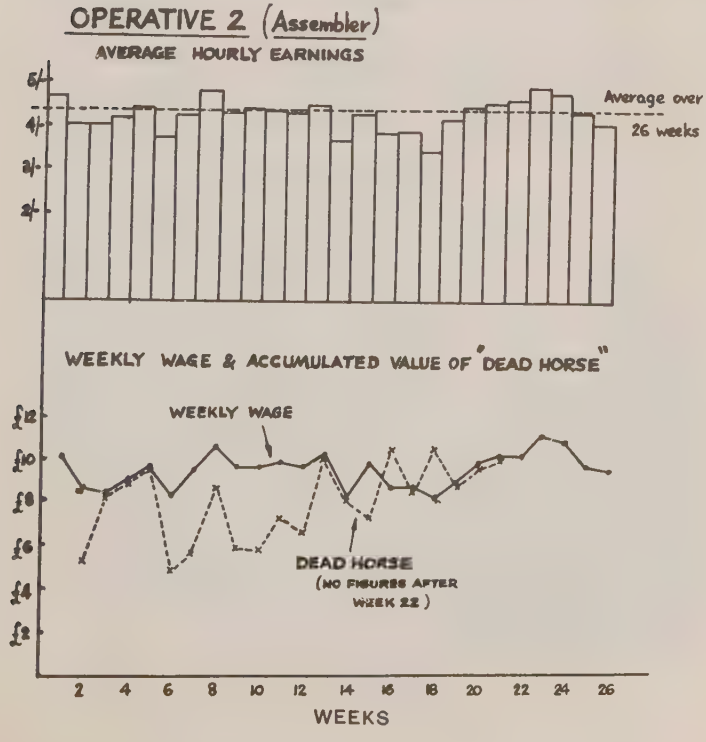
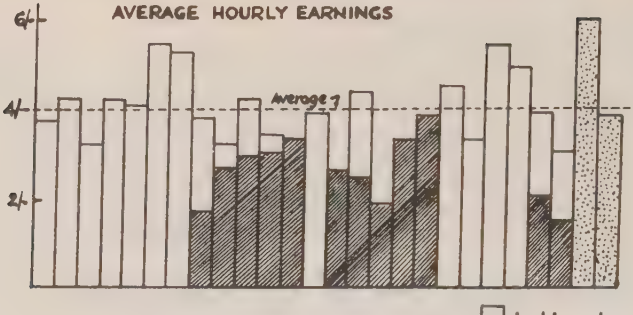
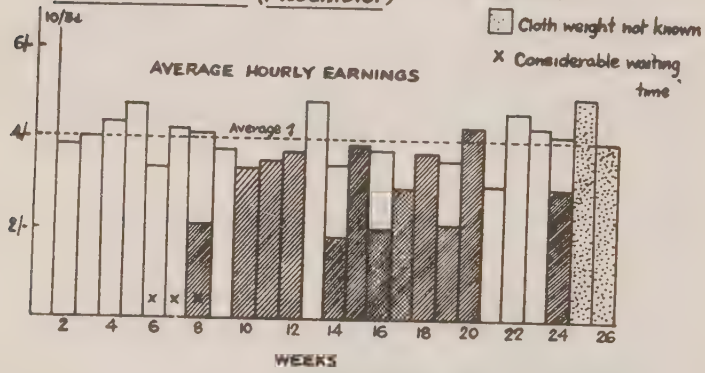

OPERATIVE 2 (Assembler)
AVERAGE HOURLY EARNINGS
5/-
4/-
3/-
2/-
Average over
26 weeks
WEEKLY WAGE & ACCUMULATED VALUE OF "DEAD HORSE"
£12
£10
£8
£6
£4
£2
WEEKLY WAGE
DEAD HORSE
(NO FIGURES AFTER WEEK 22)
2 4 6 8 10 12 14 16 18 20 22 24 26
WEEKS

FIG. 5

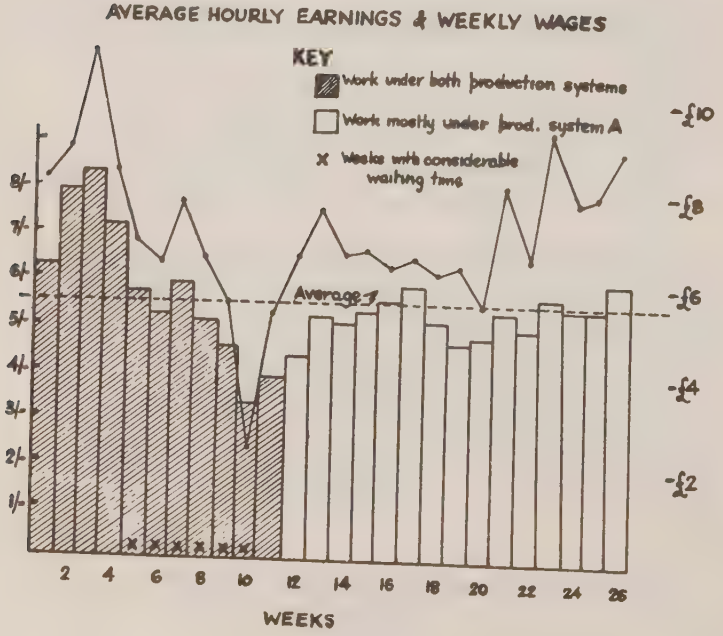
OPERATIVE 3 (Machinist)



OPERATIVE 4 (Assembler)



OPERATIVE 5 (Machinist)



Operative 1 and 2.

Fig. 4 shows average hourly earnings, weekly wage, and accumulated value of "dead horse" at the end of each week for two assemblers. The assemblers were experienced operatives, who worked usually on heavy work. They had both been with the firm many years. Neither operative experienced much waiting time. The fluctuations in wages must be traced mainly to variations in style and design of garment.

Operative 1 was very highly skilled. He made a lot of sample work and special orders. His earnings were not very high, partly due to a physical disability and partly to the fact that he was given difficult work to make. The average for the 26 week period was 3/10d. His earnings, though low, were very steady compared to most workers. The range between the highest and lowest average hourly earnings is near the lowest for the whole shop. It is 1/5d., a difference of about £3.2.4. between the weeks when the highest and lowest earnings were made, on the basis of a forty-four hour week. The relation between the weekly wage and the amount of "dead horse" for weeks four to sixteen shows clearly that "dead horse" was used to offset fluctuations in wages. This operative received somewhat special treatment from management. Though he was given very difficult work to make in some weeks, in others he was given work which was easy to make money on, thus keeping his wage steady. His average hourly earnings were low from weeks fourteen to nineteen. The styles in this period were particularly difficult. There was much navy or black work, which is slow to make because it strains the eyes. The excessively high amount of "dead horse" in week 20 was the result of Operative 1 taking a long time to make a very intricate and expensive special order garment.

The earnings of Operative 2 fluctuate more, but the range is still not very high. The average for the 26 weeks was 4/3d., and the range 1/5¾d. The low average hourly earnings in weeks two to four were the result of work on black or navy work. The cause of the low earnings from week 14 to week 20 was the same as that for Operative 1—difficult styles, and black or navy work. The graph showing the relation of "dead horse"

to the weekly wage shows that "dead horse" did more than offset weekly fluctuations. There is a rising trend through the period, probably due to the fact that trade was getting a little easier. There are several weeks, e.g. week 6, where changes in the amount of "dead horse" can be seen to be a cause of the change in the wage.

Operatives 3 *and* 4.

In fig. 5 the average hourly earnings and accumulated value of "dead horse" at the end of each week, are shown for a machinist, Operative 3, and an assembler, Operative 4. Much of the fluctuation in earnings is, as in the previous cases, due to changes in design. The major factor at work however was a change from light to heavy work. Both these workers worked only on light work when trade was busy. For weeks in which they worked on heavy work the columns representing earnings are shaded. Where a substantial amount of wages in any week was earned by work on both light and heavy work, part of the column only is shaded. The area shaded corresponds roughly to the proportion of wages earned on heavy work in that week.

Both the range in earnings and the association between high average hourly earnings and light work is more marked for Operative 3 than for Operative 4. For Operative 3 the range is 3/4d., or £7.6.8. for a forty-four hour week. The average for the 26 weeks is 4/0¼d. For Operative 4 the average is 4/0¼d., but the range is only 1/11½d. The average hourly earnings of Operative 4 were lowered in weeks 6, 7 and 8 by waiting time. Work for assemblers was short at this time and the higher earners among the assemblers, which included Operative 4, were on an informal quota. Low average hourly earnings occurred in weeks 10, 17 and 18 because these were weeks with holidays. The weekly wage was supplemented by holiday pay. This is calculated on the earnings of the previous financial year, and average hourly earnings in that year were considerably higher. Earnings of both operatives were higher in week 25. This may have been the result of the type of work made in that week, but we have no information about the type of work made after week 24.

Operative 5.

The fluctuations in earnings of Operative 5, a machinist, were greater than those for most other operatives. The range in average hourly earnings over the twenty-six weeks was 5/0½d., or £10.1.10. for a forty-four hour week. The range was nearly as big as the average hourly earnings for the twenty-six weeks taken together, which was 5/6d. The reason for these great fluctuations was excessive waiting time caused by the shortage of work on two garments which constituted nearly all the work made under Production System B. One of the two garments made under this system, a small boy's coat, paid particularly well on account of its size. Production on these two types of garment began to slacken off after week 4. By week 8 no more of one type were being made, and very little of the other. Operatives who worked mainly on these garments experienced a severe drop in wages up till week 10. The graph illustrates this for Operative 5. Holiday pay has been left out of the total wage for weeks 10, 17 and 18, and so the take-home pay was higher than is shown here.

Machinists who worked with assemblers solely or mainly under Production System A did not suffer such a drop in wages. It would have been against the traditional practice of the factory to re-allocate the work of assemblers and share it out between all machinists. However, in week 10 several machinists left, because of low earnings. They included some who had worked with assemblers. Also in week 11 several new workers, all assemblers, started work. The work of these assemblers was given to machinists who were short of work. This meant that there were fewer machinists working solely under Production System B among whom to share out that work. The quantity of work made under Production System B did not, however, begin to increase again until about week 20. These circumstances affected the earnings of Operative 5. She was working mainly under Production System B until week 12, and experienced a severe drop in wages until week 10. After this her wage recovered, but it only began to rise to the level achieved under Production System B in week 21. Average hourly earnings followed the same pattern as wages, but hourly

earnings, in contrast to wages, never reached the high level they had been under System B.

Analysis of the average hourly earnings of these five operatives brings out the importance of the effect of market factors. The main reasons given for the fluctuations are :

Operative 1 *and* 2.
Changes in design and materials.

Operatives 3 *and* 4.
Changes from light to heavy work, and changes in design.

Operative 5.
Waiting time, small size of bundles, changes in design resulting in change from Production System B to A.

Average hourly earnings in boom and slump compared.

Table VI, comparing average hourly earnings for assemblers and machinists for two periods in 1954 and 1956, serves to emphasise the effect of market factors. 1954 was a period of brisk trade, when there was more work to be done than there were operatives to do it. It stands in marked contrast to the depression of 1956. In almost all cases operatives suffered a drop in earnings in the 1956 period, and in some cases a very substantial one (32% for one worker). There were only three substantial increases. Two of these cases are foreign operatives who were learning the trade in 1954. We can give no reason for the other increase.

For assemblers the drop ranges from 1% to 24%. The relative fall in average hourly earnings was least among operatives who worked almost invariably on heavy work, as for Operatives 7 and 11. The only increases occurred among this group of Operatives 10, 12 and 15. The drop was most for those who worked almost always on light work, as for Operatives 2, 3, 4 and 5. These operatives were most frequently called "killers" by other workers in the factory. It is probable that the quality of their work was below that of most other operatives. It was more difficult to get poor quality work past the inspectors and the manager in slack periods, and this explains the large drop in earnings. An exception to this is worker 6. She usually

TABLE VI

COMPARISON OF AVERAGE HOURLY EARNINGS OF ASSEMBLERS AND MACHINISTS FOR THE 1954 AND 1956 PERIOD

Number of assembler or machinist	*Assemblers*		Differences between 1954 & 1956 as a % of 1954	*Machinists*		Differences between 1954 & 1956 as a % of 1954
	1954	1956		1954	1956	
1	5/6	4/9	−14	8/6	7/6¾	−11
2	5/1½	4/4¾	−14	8/2¾	7/4¾	−10
3	5/0½	3/9¾	−24	6/5¾	5/5½	−15
4	4/11½	4/0¼	−19	5/11	4/0¼	−32
5	4/11	4/2¾	−14	5/11	4/7¼	−21
6	4/9¼	4/6½	− 5	5/7¼	4/6¼	−19
7	4/3½	4/3	− 1	5/6½	4/4¾	−21
8	3/10½	3/7¼	− 7	5/5½	4/6¾	−16
9	3/10¼	3/5½	−10	4/11½	4/4¼	−12
10	3/9½	3/7¼	− 5	4/10	4/2	−14
11	3/9¼	3/10	+ 2	4/9	3/9¼	−21
12	3/6	3/8½	+ 6	4/7¾	4/4¾	− 5
13	3/5¾	2/11½	−15	4/7¾	3/2	−32
14	3/5	3/1	−10	4/7¾	3/10½	−17
15	3/3¾	4/–	+21	3/11	3/5¼	−12
16	3/3	3/0½	− 6	3/9¾	3/4¾	−11
17	3/3	2/7¼	−20	3/8¼	3/2	−14
18	2/11¾	2/9¾	− 7	3/5	2/6½	−26
19	2/11¼	2/9	− 6	3/4¼	4/11½	+47
20	2/11¼	2/10	− 4	2/9½	3/1¾	+13

Note :

Average hourly earnings are based on 20 weeks in the 1954 period and 26 in the 1956 period, except in the following cases :—

1954. For assemblers 5, 7, 8, 17, 18, 19 and 20 figures are based on 18, 19, 9, 16, 19, 19 and 19 weeks respectively.

For machinists 3, 5, 7, 8, 10, 12, 15, 17 and 19 figures are based on 11, 9, 7, 17, 19, 8, 19, 17 and 17 weeks respectively.

1956. For assemblers 10, 14, 17 and 20 figures are based on 25, 24, 16 and 25 weeks respectively.
For machinists 5, 6, 7, 10, 11, 12, 14, 15, 17 and 18 figures are based on 10, 11, 8, 22, 25, 10, 23, 11, 4 and 11 weeks respectively.

Only operatives who worked at the factory during both periods are included in this table.

worked on light work and was known as a "killer." Yet her earnings dropped by only 5%.

The drop in earnings for machinists is greater than that for assemblers, ranging from 5% to 32%. This is most likely due to the shortage of work made under Production System B. It is not possible to compare the drop in earnings of different operatives because in the 1956 period some operatives stayed around the factory when there was no work, while others went home.

We have no field work on the 1954 period.

It has not been possible to illustrate the effect on earnings of such things as absence, and sickness. Their effect is masked by "dead horse" in a way which is not possible in the case of factors acting over a longer period. Though they cannot be related to specific fluctuations in average hourly earnings calculated over the period of the week, this does not mean that they were unimportant in their effect. To the workers they often seemed more important than other factors. This was not only because they were more immediate, but also because the effect of such things as shortage of work came to be taken for granted by operatives after many years of hard experience of seasonal slumps in trade, not to mention the long slump period which affected the trade throughout the 1930's.

Summary.

This account of the operation of a piecework incentive scheme shows clearly that the workers did not receive equal cash reward for equal effort. It is true that we have no objective measurement of effort. We rely on our own observations and on what operatives have told us in estimating the effect of the variation in effort expended compared with the effect of other factors noted, for instance waiting time. But we have been able to show that there are factors other than effort expended which help to determine average hourly earnings. And we have related these other factors to fluctuations in average hourly earnings for a selected number of operatives.

In Factory III, the "cash content" of an hour's work varied tremendously, not only from boom to slump, but also from one week to another. At any one time, because of the

anomalies in the price list, while one operative had work which was easy to earn money on, another had work which it was difficult to earn money on. Such situations created a great deal of jealousy among operatives. In this factory unequal cash reward for equal effort led operatives to attempt individually to exercise control over earnings by means of the modifying factors discussed above. The fact that operatives tried individually meant that they almost invariably failed, for management could play one off against the other.

We conclude that this piecework incentive scheme was inadequate because it failed to take into account important factors influencing the operatives' ability to earn. It would seem from the nature of these factors that they would be best dealt with at the level of the workshop, and that some form of workshop negotiation about piecework prices is needed.

Conclusion.

Our analysis of the operation of three incentive schemes revealed the existence of "built in" anomalies. In Factory I and Factory II the method of fixing job prices and times yielded "tight" and "loose" rates. In Factory III the fixing of the relative prices of items in the Price List was the main cause of "good" and "bad" work. Even if there had been no hold-ups in the flow of work, no sickness, no absence, and a steady flow of orders, the workers' earnings would still have been subject to fluctuation because of anomalies in the job prices.

But incentive schemes do not operate in Utopia. They operate in the context of production for a competitive market. Inevitably, hold-ups occur. These may occur in a regular and foreseeable pattern, in which case they can be systematically dealt with. But there are always haphazard and irregular disturbances brought about by human error, sickness, absence, and the like ; and these affect the capacity of the worker to earn.[1] And for such disturbances cash compensation is difficult to arrange, even if it is accepted that all risks of this kind

[1]When considered statistically over a large universe such occurrences as these may exhibit regularities. At workshop level their incidence appears haphazard.

should be borne by management—and there is often disagreement on this point.[1] But even where attempts are made to compensate the worker for the effects of such factors they are likely to fall short of entire success, as our evidence has indicated. So the effect of the productive context in which an incentive scheme operates is often to emphasise the anomalies which are built into the scheme itself.

Market conditions, as we have shown, also exert a disturbing influence on earnings. This effect was more marked in the clothing workshops than in the engineering workshop. But even in the engineering workshop the variety of work, a feature of the market situation confronting the firm, had its effect on the allocation of work and hence on the earnings of the operatives.

In the two clothing factories attempts made by workers to control those factors in the working environment which affected the amount and stability of their earnings were limited and not very effective. We also note that such attempts were not made collectively but by individuals. In the engineering factory the workers had some success in their efforts collectively to exercise control. The manipulation of "booking-in" to stabilise earnings by offsetting the effects of waiting time, and as a defence against possible rate-cutting, was regarded by most of the operatives as being slightly unethical, and was eschewed by some. But in a situation where "straight" booking seemed to place them at a disadavantage, the "fiddle," as they called it, seemed justifiable to most of the workers.

Our analysis would seem to support the view that, to make incentive systems work effectively is partly a problem for production engineers, and partly a social engineering or "human relations" problem. It could be argued that to devise more accurate techniques of job rating would remove "built-in" anomalies, and to perfect techniques for planning and controlling work flow would eliminate hold-ups and the incidence of waiting time. It would seem that the solution of the production engineering problems would do much to encourage greater

[1]We know of incentive schemes designed so that management bears all risks of this kind, and which guarantee stability of earnings, but our limited evidence does not suggest that such schemes are widely used.

co-operation between managers and workers, since it would remove irritations and the need for "fiddles." But there would still remain the problem of workers' suspicions of the motives of management, for example, fear that management will cut job rates if earnings are high. To remove these suspicions is not a production engineering problem but one of devising management techniques which will generate confidence and goodwill, and a feeling of security.

We would agree that much could be done to improve the operation of piecework incentive schemes by an improvement in production engineering and management techniques. But we would argue that however technically perfect a system of ratefixing may be, judgments about what constitutes a "fair" effort and a "fair" reward are still involved, and room for dispute remains. We are not convinced by arguments that methods of ratefixing can be devised which eliminate evaluation.[1] Neither are we convinced that detailed production control to eliminate hold-ups can always be entirely effective. In addition to our own evidence there is much other evidence to show that in those factories where much attention has been given to devising foolproof methods of control, problems are still encountered which are as difficult to deal with as those encountered under "old-fashioned" methods.[2] One reason for this is that from the workers' point of view, tight management control may well be seen as another obstacle in the way of efforts to make the job pay. This is quite apart from resentment generated by constraints upon workers' behaviour which are implicit in such systems of control.

The evidence adduced in this paper, taken together with other evidence, and our own experiences on the shop floor, lead us to conclude :

[1]For a more detailed discussion of this point see :
Friedmann, George, *Industrial Society*, The Free Press, Glencoe, Ill. (1955).
Lupton, Tom, "A Sociologist looks at Work Study," *Work Study and Industrial Engineering*, Vol. I, No. 2, Feb., 1957.

[2]See, for example :
Roy, Donald, "Efficiency and 'the Fix' ", *Amer. J. Sociol.*, IX, No. 3, Nov., 1954. Roy describes the success of the workers in a piecework shop "in circumventing each of a series of "new rules" and "new systems" introduced by management."

1. Methods of job rating are still crude. In this we are in agreement with Barkin, a Trade Union observer of American industry, who observes :

   "Current techniques for measuring human application are at best only makeshifts, lacking a scientific basis, and providing no objective basis whatever for reconciling employer and Trade Union interests." [1]

   If this is so, then it would seem to follow that there remains much scope for "on the spot" bargaining to take account of differing judgments. Since the worker is likely to be in possession of knowledge of the immediate factors affecting his capacity to earn, to involve him fully in the rate-fixing process seems to us a reasonable and sensible course.
2. Since there are factors in each workshop which are not susceptible to close control because of uncertainty and unpredictability, it would seem wise to plan flexibility into a system so that it can the more easily adapt to unforeseen changes. Our analysis has shown, for example, how great can be the effects of changes in the market upon earnings, *via* waiting time, changes in product design, and so on. To attempt to plan rigidly for such eventualities when knowledge is incomplete and unreliable is to invite the workshop opinion that management does not know what it is doing. To admit that knowledge is incomplete, and to discuss how commonly shared hazards are to be handled seems to us to be a more reasonable course.
3. Our analysis seems to offer at least a partial explanation of the state of affairs described by Barkin :

   "Rather than unleashing a stream of productivity increases, incentive wage systems have in some places provoked interest in securing bonus without raising output. Debilitated quality, sabotaged morale, and worker resistances have been the only real products of the plan." [2]

[1]Barkin, S., "Concepts in the measurement of Human Application," *Ind. and Lab. Rels. Review*, Vol. 7, No. 1, Oct., 1953.

[2]Barkin, S., "The Bench-Mark Approach to Production Standards," *Ind. and Lab. Relations Review*, Vol. 10, No. 2, Jan., 1957.

If an operative, or a group of operatives, observes that the relationship between the effort they exert, and the cash reward they receive, is affected by factors which are beyond their control, it seems natural that they will attempt to redress the balance by operating such controls as are open to them. Depending on circumstances one may find attempts to "fiddle" on booking, on quality, or on work allocation. The extent to which these efforts are successful will depend upon factors peculiar to each workshop; such as the strength of Trade Union organisation, the degree of machine pacing of the process, and the like.

We consider that the worker's desire to control some of the factors which affect his earnings is a legitimate one. If the setting of rates and the design of new productive methods were to be a fully joint activity, much of the energy which workers divert to devising controls to defeat the ends of management could be directed into constructive channels for mutual benefit.

Tom Lupton
Sheila Cunnison

University of Manchester.

Factor Endowments, International Trade and Factor Prices

In the past few years, there has been a revival of interest in the Heckscher-Ohlin model of international trade, and a closer scrutiny of two propositions associated with its analysis of the principle of comparative costs.[1] The first is that the cause of international trade is to be found largely in differences between the factor-endowments of different countries; the second that the effect of international trade is to tend to equalize factor prices as between countries, thus serving to some extent as a substitute for mobility of factors. The result, very briefly, has been to show that neither proposition is generally true, the validity of both depending on certain factual assumptions about either the nature of technology or the range of variation of factor endowments which are additional to, and much more restrictive than, the assumptions of the Heckscher-Ohlin model itself.

The purpose of this article is to survey the related problems of the relation between factor-endowment and international trade and the effect of trade on relative factor prices, with the aim of clarifying the nature and simplifying the explanation of some of the results of recent theoretical research. The main instrument employed to this end is a diagrammatic representation of the technological side of the economy, developed from one originated by Mr. R. F. Harrod.[2]

[1]See in particular:
Abba P. Lerner, "Factor Prices and International Trade," *Economica*, N.S., XIX, No. 1, Feb., 1952, 1-15;
I. F. Pearce, "A Note on Mr. Lerner's Paper," *loc. cit.*, 16-18;
S. F. James and I. F. Pearce, "The Factor Price Equalisation Myth," *Review of Economic Studies*, XIX (2), No. 49, 1951-52, 111-20;
Romney Robinson, "Factor Proportions and Comparative Advantage: Part I," *Q.J.E.*, LXX, No. 2, May, 1956, 169-192;
R. W. Jones, "Factor Proportions and the Heckscher-Ohlin Theorem," *Review of Economic Studies*, XXIV (1), No. 63, 1956-57, 1-10;
also
Lionel McKenzie, "Equality of Factor Prices in World Trade," *Econometrica*, 23, No. 3, July, 1955, 239-57, which handles the same problems by linear programming methods.

[2]Presented at the conference of the Association of University Teachers of Economics, January, 1957.

In common with other models of international trade, it is assumed, on the consumption side, that tastes and the distribution of the means of satisfying wants (property ownership, or claims on the social dividend) are given ; and that, on the production side, technology and the supply of factors of production in each country are given (the latter implying that factors are immobile between countries). Production functions and factors are assumed to be identical in all countries ; the outputs of goods are assumed to depend only on inputs of factors into the production processes for those goods, and factors are assumed to be indifferent between uses. Further, production is assumed to be subject to constant returns to scale, so that the marginal productivities of factors depend only on the ratios in which they are used. Finally, perfect competition and the absence of trade barriers (tariffs and transport costs) are assumed. The argument is simplified still still further by assuming, initially, the existence of only two countries, I and II, two goods, X and Y, and two factors of production, labour and capital,[1] the available quantities of the factors being given independently of their prices.

The given technical possibilities of production, summarized in the production functions, imply a definite relationship between the optimum capital:labour ratios in the two industries, relative factor prices, and relative costs of production of commodities. This relationship is represented in Fig. 1. For any given relative price of labour in terms of capital (such as w_I) there will be an optimum ratio of capital to labour in the production of each good ($r_{x.e}$ in X and $r_{y.e}$ in Y) which equates marginal productivities to factor prices, and a relative cost of X in terms of Y (c_I) which embodies the ratio of their costs of production at the given factor prices with the associated capital:labour ratios.

An increase in the relative price of labour above the given level will make it profitable to substitute capital for labour in both industries, thus raising the optimum capital:labour ratio

[1]The simplicity of these assumptions is a virtue rather than a defect, in an argument whose purpose is to show how little can be said.

in both :[1] it will also increase the relative cost of the labour-intensive good (X in the neighbourhood of w_1), since its costs will be raised more than those of the capital-intensive good by the increase in the relative price of labour. Further increases in the relative price of labour will continue to raise the optimum capital:labour ratio in both industries ; but either one of two possibilities may be the effect on the relative cost ratio, depending on the relative ease of substituting capital for labour in the two lines of production. The first, and simpler, occurs when one commodity remains labour-intensive and the other capital-intensive, whatever the relative factor price. In this case, one commodity can be definitely identified as labour-intensive and the other as capital-intensive ; and the relative cost of the labour-intensive good will continue to rise as the relative price of labour rises, so that the latter can be deduced from the former. The second occurs when, owing to greater facility in substituting capital for labour in the initially labour-intensive good, the difference in capital-intensity between the two goods narrows and eventually reverses itself, so that the labour-intensive good becomes capital-intensive and *vice versa* ; such a reversal of factor-intensities may occur more than once (as illustrated in Fig. 1), as variation in relative factor prices induces variation in the capital-intensity of production of both goods. In this case, a commodity can only be identified as labour-intensive or capital-intensive with reference to a range of relative factor-prices or (what is the same thing) a range of capital-intensities ; and the relative cost of a commodity will alternately rise and fall as the relative price of labour rises, as that commodity varies from being labour-intensive to being capital-intensive. Consequently more than one factor price ratio may correspond to a given commodity cost ratio, and the factor price ratio cannot be deduced from the commodity cost

[1]The shapes of the curves relating optimal capital:labour ratios to the relative price of labour assume increasing difficulty in substituting capital for labour as the capital-output ratio rises and continuous substitutability of labour for capital in response to changing relative factor prices. On linear programming assumptions the curves would be step functions, the steps representing discontinuous changes in capital intensity. At their left extremities, the curves will cut the *Ow* axis if labour can entirely replace capital in production, and the *Or* axis if it cannot.

ratio alone. The distinction between the two cases, which turns on the facts of technology, is fundamental to what follows.

The analysis so far has been concerned with the relationships implicit in the given technological possibilities of production. But only a limited range of the techniques available can actually be used efficiently by a particular economy with a given factor endowment ; and the possible range of relative factor prices and relative commodity costs is correspondingly limited. The factor endowment of the economy sets an overall capital:labour ratio, to which the capital:labour ratios in the two industries, weighted by the proportions of the total labour force employed, must average out. At the extremes, the economy's resources may be used entirely in the capital-intensive industry or entirely in the labour-intensive industry ; and the relative prices of factors, and the relative costs of commodities, must lie within the limits set by these two extremes.[1]

The restrictions imposed by the factor endowment on the techniques the economy can employ and on the possible variation of factor prices and relative costs of production are illustrated in Fig. 1, where r_I represents the economy's overall capital:labour ratio. If all resources are used in the production of Y, the relative price of labour will be w_1, if all are used in the production of X the relative labour price will be w_2 ; the corresponding relative costs of commodity X are c_1 and c_2. As w rises from w_1 to w_2, the capital:labour ratio in Y rises from r_I to $r_{y.o}$ and the ratio in X from $r_{x.o}$ to r_I ; the increases in these ratios are reconciled with the constancy of the overall capital:labour ratio by a shifting of resources from production of Y to production of X, which frees the capital required for the increases in the ratios. It can easily be shown that, for any relative factor price, the proportions of the total labour supply employed in one of the industries is equal to the ratio of the difference between the other industry's capital:labour ratio and

[1]At the extremes, the country is completely specialized on one or other commodity ; in a closed economy, there would therefore be no exchange ratio between commodities, though the non-produced one might have a "virtual" price (equal to or less than its cost of production) ; in an open economy, the price of the non-produced good would be an international one (equal to or less than the domestic cost of production).

the overall ratio, to the difference between the capital:labour ratios in the two industries ; thus, for example, the proportion of the labour force employed in producing X at the relative labour price w_I is $(r_{y.e} - r_I)/(r_{y.e} - r_{x.e})$.

As can readily be seen from the diagram, the range of variation of relative factor prices and commodity costs possible in the economy depends broadly on two technological factors : the spread between the optimum capital:labour ratios in the two industries (reflected in the vertical distance between the r_y and r_x curves), and the difficulty of substituting capital for labour (reflected in the gentleness of the slopes of the two curves). The greater the difference between optimum factor-intensities and the greater the difficulty of substituting capital for labour, the greater the range of variation of factor prices and commodity costs possible between the extremes of specialization on one or other product. If, on the other hand, capital:labour ratios were the same in the two industries or factors were perfect substitutes, there would be no possibility of variation of relative factor prices and commodity costs.

Before proceeding, it may be as well to notice that the assumption that the quantities of factors available are fixed, can be relaxed without trouble. Dependence of factor supplies on their relative prices can be readily introduced by making the overall capital:labour ratio depend on the relative price of labour ; in the normal case the curve representing the relationship will be downward-sloping, reflecting the reduction in the supply of capital and increase in the supply of labour as the price of the former falls relatively to the latter. The effect (in the normal case) would be to restrict the range of possible variation more than would otherwise be the case ; in the extreme case of perfect elasticity of factor supply, only one relative factor price ratio and commodity cost ratio would be possible.

Within the range of variation permitted by technology and factor endowment, the factor price and commodity cost ratios will be determined by the forces of demand. For a closed economy, this determination is most readily conceived as follows : to any given relative price of labour will correspond a certain pattern of production, a certain distribution of

income, and a certain price ratio between the commodities at which the recipients of the income will be just willing to consume the quantities of both goods produced at the given price of labour. As the relative price of labour rises and resources are shifted out of producing the capital-intensive commodity into producing the labour-intensive commodity, the equilibrium demand price of the latter is very likely to fall,[1] while its relative cost is rising. The equilibrium levels of commodity prices (costs) and factor prices will be determined by the condition that the relative costs of production should be just equal to the relative prices at which the outputs of the two commodities will be absorbed.[2]

The next stage of the argument is to introduce the possibility of international trade and examine its effects. This is done in Fig. 2, which reproduces Fig. 1 with the deletion of certain parts of it no longer necessary to the argument. In the diagram, the country previously analysed ("country I") is represented by its factor endowment ratio r_I, and its closed-economy equilibrium factor price and comparative cost ratios w_I and c_I. The possibility of trade is introduced *via* another country ("country II") similarly represented on the diagram by its factor price and cost ratios w_{II} and c_{II}, with a different closed-economy equilibrium comparative cost ratio. It is assumed in all cases that country II is endowed with a higher ratio of capital to labour than country I.

So long as the closed-economy equilibrium comparative cost ratios of the two countries are different, international trade will be profitable, and each country will export the good in which it has a comparative cost advantage (as measured by

[1]The exceptional case is that in which the reduction in the supply of the capital-intensive good is offset by a reduction in demand because the owners of labour have a stronger preference for the labour-intensive good than the owners of capital. It makes no fundamental difference to the analysis unless the effect is so strong that the equilibrium demand price of the labour-intensive good rises more rapidly than its cost of production, a possibility of instability which will be ignored here.

[2]The determination of full general equilibrium could be represented by drawing in the bottom half of Fig. 1 a curve expressing the equilibrium price of X as a function of the relative price of labour (and, implicitly, of the factor endowment), whose intersection with the relative cost curve would determine the equilibrium of the system.

equilibrium commodity price ratios in the absence of trade). In the assumed absence of trade barriers, the effect of trade will be to equate the price ratios in the two economies at a level (c_e in the diagram) somewhere between the two closed-economy comparative cost ratios, this level being determined by the demand and supply conditions in the two countries taken together.

At the new international price ratio, both countries may continue to produce both commodities, or one may specialize completely on the production of the commodity in which it has the comparative advantage, or both may specialize. Which will be the case is governed by two principles—that the new price ratio will lie between the two closed-economy price ratios, and that a country will not specialize at a price ratio between the extremes of its comparative cost range—and is determined by the relations between four cost ratios—the maximum cost ratio in the country with the lower maximum, the minimum cost ratio in the country with the higher minimum, and the two closed-economy equilibrium price ratios. The following four propositions can be deduced from the two principles stated above ; for diagrammatic simplicity, they are illustrated only for cases in which the two countries' capital:output ratios lie in the same range, so that there is a one-to-one relation between cost ratios and factor price ratios.

(*i*) If the minimum cost ratio in the country with the higher minimum exceeds the maximum cost ratio in the other country, so that the comparative cost ranges of the two countries do not overlap, one country at least must specialize completely. This would be the case if country II's endowment ratio coincided with r_a in Fig. 2.

(*ii*) If the cost ratio ranges of the two countries overlap, at least one of the countries will not specialize. This would be the case if country II's endowment ratio were either r_b or r_{II} in Fig. 2.

(*iii*) If one country's cost ratio range includes the other's closed-economy equilibrium price ratio, the former country will not specialize. For example, if country II's endowment ratio were r_b and its closed-economy

equilibrium labour price w_b, the labour price ratio corresponding to the new price must fall between w_b and w_I, so that country I would produce both goods.

(*iv*) If each country's closed-economy equilibrium price ratio falls within the other's cost ratio range, neither will specialize completely. For example, with country II's endowment ratio r_{II} and closely-economy equilibrium labour price w_{II}, the labour price ratio corresponding to the new equilibrium price must fall between w_I and and w_{II}, in which range both countries would produce both goods.

In the explanation of the causes of international trade presented above, comparative advantage was measured by equilibrium prices in the absence of trade. Such an explanation, amounting to the statement that those goods are traded which it is profitable to trade, is little more than a tautology. The next problem requiring analysis, and the one which the apparatus built up earlier in this article is designed to elucidate, is the relation between comparative advantage so defined and differences in factor supplies. Closely related to this problem is the problem of the effect of international trade on the prices of factors, which can be dealt with by modifications of the same argument.

To begin with, it is necessary to notice a difficulty : relative scarcity or abundance of a factor may be defined in either of two ways—relative cheapness, a matter of price, and relatively high endowment ratio, a matter of physical quantity. These two definitions may not be synonymous, and it is necessary to investigate the relationship of differences in factor supplies on both definitions to differences in comparative advantage.

The investigation is most conveniently divided into two parts, corresponding to the two cases distinguished in the discussion of the technological data above.

The first case is that in which there is no reversal of the factor-intensities of commodities in the range of capital:labour ratios between the endowment ratios of the two countries. This situation (exemplified by the combination of r_I with r_a,r_b, or r_{II} in Fig. 2) may be the consequence either of

technological conditions which exclude the possibility of such a reversal, or (as in the diagram) of sufficiently similar endowment ratios in the two countries. In this case, owing to the one-to-one relation between comparative cost and factor price ratios, a relatively lower commodity price in one country necessarily implies a relatively lower price there of the factor used relatively intensively in the production of that commodity. Hence comparative advantage is necessarily associated with factor abundance, and disadvantage with factor scarcity, when abundance and scarcity are defined in terms of cheapness or dearness of the factor concerned in one country as compared with the other.

But scarcity and abundance in this sense are not necessarily the same as scarcity and abundance as measured by factor endowment ratios, since factor prices depend on demand as well as on supply, and the influence of demand may outweigh the influence of supply.

Thus a country with relatively abundant labour, judged by its endowment ratio, may be a country with a relatively high price of labour, and hence export the capital-intensive commodity. This possibility is illustrated in Fig. 2 by the combination of r_I c_I with r_{II} c_{II} : country I is at a comparative disadvantage in the production of X, the labour-intensive commodity ($c_I > c_{II}$), labour being more expensive there than in country II ($w_I > w_{II}$) ; but this is quite consistent with country I having a relatively more abundant endowment of labour, and less abundant endowment of capital, than country II ($r_{II} > r_I$).

For this possibility to occur, two conditions must be fulfilled. The first is that each country's comparative cost range must overlap the other's closed-economy equilibrium price ratio, which entails (proposition (*iv*) above) that international trade will lead neither country to specialize completely. The second, and more important, is that each country must have a strong preference for the good which uses intensively

the factor with which the country is relatively abundantly endowed.[1]

The effect of trade on relative factor prices in this case is a comparative simple matter : owing to the one-to-one relationship between commodity cost ratios and relative labour prices, the equalisation of commodity prices through trade will tend to equalize relative factor prices.[2] If both countries continue to produce both goods under international trade, equalisation will necessarily be complete. (In the case shown in Fig. 2, the relative price of labour in both countries will be w_e, the price corresponding to the common commodity price ratio c_e).[3]

The second case is that in which the endowment ratios of the two countries are separated by one or more reversals of factor-intensities in the production of two goods. In this case, there is no possibility of the country with the higher capital: labour endowment ratio having a lower relative price of labour, since by assumption the ranges of relative factor prices possible in the two countries do not overlap : factor scarcity and abundance can therefore be measured indifferently by endowment ratios or relative factor prices. But because of the intervening factor-intensity reversal(s) and the associated reversal(s) of direction of the effect of a relative factor price change on the commodity cost ratio, it is not possible to relate comparative cost advantage in the production of a particular good to relative abundance (in either sense) of the factor utilised relatively intensively in producing that good—or even, in some cases, to identify goods as relatively labour- or capital-intensive in the

[1]Using the ratio measure of the proportions of labour employed in the two industries referred to above, it can easily be shown that at the pre-trade country I price of labour country II would have a smaller proportion of its labour force engaged in the production of the labour-intensive good than country I ; at the lower pre-trade country II price of labour, the proportion must be still smaller. At the common price-ratio established by trade, country II must devote a larger proportion of its resources to the capital-intensive industry than will country I.

[2]Since production functions are identical in the two countries, equalisation of relative factor prices also means equalization of absolute factor prices (marginal productivities).

[3]If the equilibrium price ratio lies outside a country's range of commodity cost ratios, its relative factor price will be that corresponding to the cost ratio in the range nearest to the equilibrium price ratio.

world as a whole. The commodity in which the labour-abundant country has a comparative advantage may be either labour-intensive or capital-intensive in that country, and either capital-intensive or labour-intensive in the capital-abundant country. And since the effect of trade is to raise the relative price of the commodity in which the country has a comparative advantage, and therefore the relative price of the factor used relatively more in the production of that commodity, its effect on relative factor prices may be to move them in the same direction in the two countries, towards each other, or away from each other. Moreover, in equilibrium both countries may continue to produce both goods, even though factor prices are different in the two countries.

Two sub-cases may be distinguished for analysis of the effect of trade on factor prices, according to whether the factor-endowment ratios of the two countries are separated by an odd or an even number of reversals of factor-intensities.

If the number of reversals is odd, a commodity which is labour-intensive in one country will be capital intensive in the other, so that the export commodities of the two countries must have the same relative factor-intensity in the country of origin, being either both capital-intensive or both labour-intensive. The effect of trade will be to move factor prices in the same direction in both countries, and the difference between them may either narrow or widen according to circumstances. If the relatively labour-abundant country has a comparative advantage in its labour-intensive commodity, trade will raise the price of labour in both countries ; conversely, if its comparative advantage lies in its capital-intensive commodity, trade will lower the price of labour in both countries. The latter case is illustrated by the combination $r_I\ c_I$, $r'_{II}\ c_{II}$ in Fig. 2, where country I, the relatively labour-abundant country, has a comparative advantage in its capital-intensive commodity Y : the effect of trade is to lower the relative price of labour in both countries, from w_I to w_e in country I and from w_{II} to w'_e in country II. In the particular case shown, both countries continue to produce both commodities, though their relative factor prices are different.

If the number of intervening reversals of factor-intensities is even, a commodity is either labour-intensive or capital-intensive in both countries. The effect of trade is therefore to move factor prices in the two countries in opposite directions ; but the movement may be convergent or divergent. If the labour-abundant country has a comparative advantage in the production of the labour-intensive commodity, factor prices in the two countries will move towards each other ; if it has a comparative advantage in the production of its capital-intensive commodity, factor prices will move away from each other. This latter case is illustrated by the combination $r_I\ c_I$, $r''_{II}\ c_{II}$ in Fig. 2, where country I has a comparative advantage in its capital-intensive commodity Y and the effect of the equalization of commodity prices through trade is to lower the relative price of labour in country I from w_I to w_e and raise it in country II from w''_{II} to w''_e. In the case shown, both countries continue to produce both goods though with different relative factor prices.

The conclusions to be drawn from the foregoing analysis with respect to the influence of differences in factor supplies on the pattern of international trade, and of trade on relative factor prices, may now be summarized and illustrated by reference to the three cases shown in Fig. 2. First, it is not necessarily true that a country will export the commodity which uses relatively intensively the factor with which the country is relatively heavily endowed ; in all three cases shown, the country endowed with the lower [1] ratio of capital to labour

[1] A corollary of this result is that the so-called "Leontief scarce-factor paradox"—that the U.S.A. is a labour-abundant, capital-scarce country, deduced from the statistical finding that U.S. exports are more labour-intensive and less capital-intensive than U.S. import-competing products—rests on a theoretically invalid assumption since the nature of a country's factor endowment, relative to that of the rest of the world, cannot be inferred from the relative factor-intensities of exports and import substitutes in its domestic production. See W. Leontief, "Domestic Production and Foreign Trade ; The American Capital Position Re-Examined," *Proceedings* of the American Philosophical Society, 97 (Sept., 1953), reprinted in *Economia Internazionale* VII (1954) ; *ibid* "Factor Proportions and the Structure of American Trade : Further Theoretical and Empirical Analysis," *Review of Economics and Statistics* XXXVIII, No. 4, Nov., 1956, 386-407. To the list of critics in Footnote 2 of the latter should be added Jones, *loc. cit.* The theoretical analysis is not, of course, the only questionable feature of Leontief's argument.

exports the capital-intensive good. Second, it is not necessarily true that a country will export the commodity which uses relatively more of the factor which would be relatively cheaper in the absence of trade. This proposition will be valid only if, as a consequence either of the nature of the available technology or of the endowment of countries with factors in not too dissimilar proportions, the relative factor-intensities of goods do not reverse themselves as the capital:labour ratio varies between the endowment ratios of the two countries—the condition required to exclude the possibility of cases as r_I, r'_{II} and r_I, r''_{II} in Fig. 2. Thirdly, the proposition that trade will tend to equalize relative factor prices, and will in fact do so if both countries continue to produce both goods, is valid only on the same condition ; otherwise, trade may have any effect on factor prices, and production of both goods in both countries is consistent with widely different factor prices.[1] Consequently, it is only on the assumption that this condition is fulfilled that it can be maintained that trade is a partial or complete substitute for factor mobility, and free trade a substitute for removal of restrictions on factor movements. Thus the conclusions of the Heckscher-Ohlin model depend not only on the assumption of competition, absence of trade barriers, constant returns to scale, and so forth, but also on an empirical assumption about the nature of technology or the degree of variation in the factor endowments of countries.

In conclusion, something must be said about the extension of the technique of analysis beyond the simple two-country, two-commodity, two-factor case.

Analysis of cases involving more than two countries requires merely the representation of the additional countries by their endowment ratios and pre-trade comparative cost ratios. Similarly, additional goods can be represented by additional curves showing the relation between the optimal capital:labour

[1]Compare P. A. Samuelson, "International Trade and the Equalisation of Factor Prices." *Economic Journal*, LVIII, No. 230, June, 1948, 163-84, and *ibid*, "International Factor Price Equalisation Once Again," *loc. cit.*, LIX, No. 234, June, 1949, 181-97. The condition required for equalisation, that factor-intensities do not reverse between the endowment ratios, is assumed implicitly in the first of these articles, and stated explicitly in the second. See James and Pearce, *loc. cit.*

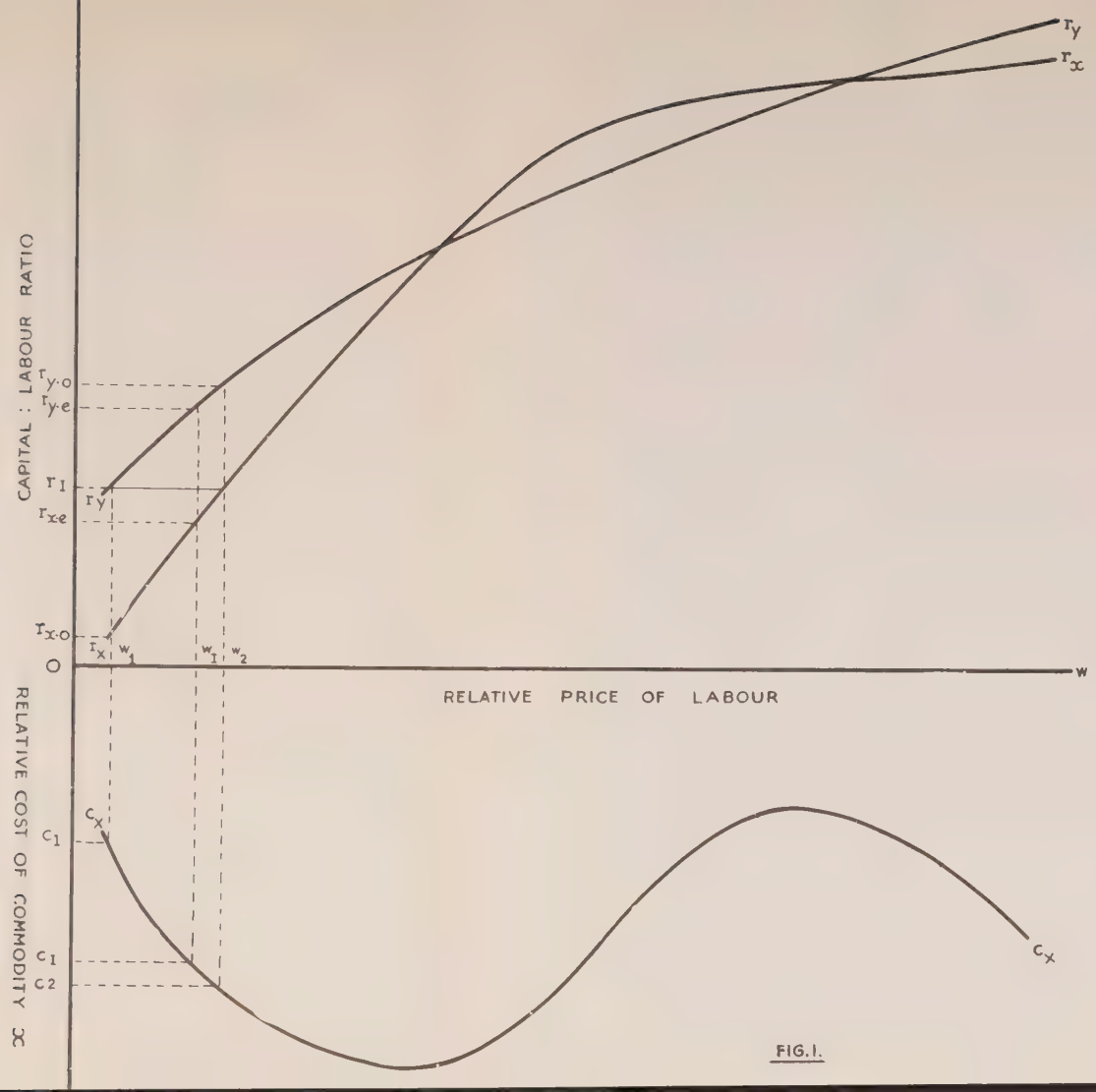

r_y
r_x
CAPITAL : LABOUR RATIO
$r_{y\cdot o}$
$r_{y\cdot e}$
r_1
r_y
$r_{x\cdot e}$
$r_{x\cdot o}$
r_x
w_1
w_1
w_2
O
w
RELATIVE PRICE OF LABOUR
RELATIVE COST OF COMMODITY X
c_x
c_1
c_1
c_2
c_x

FIG. I.

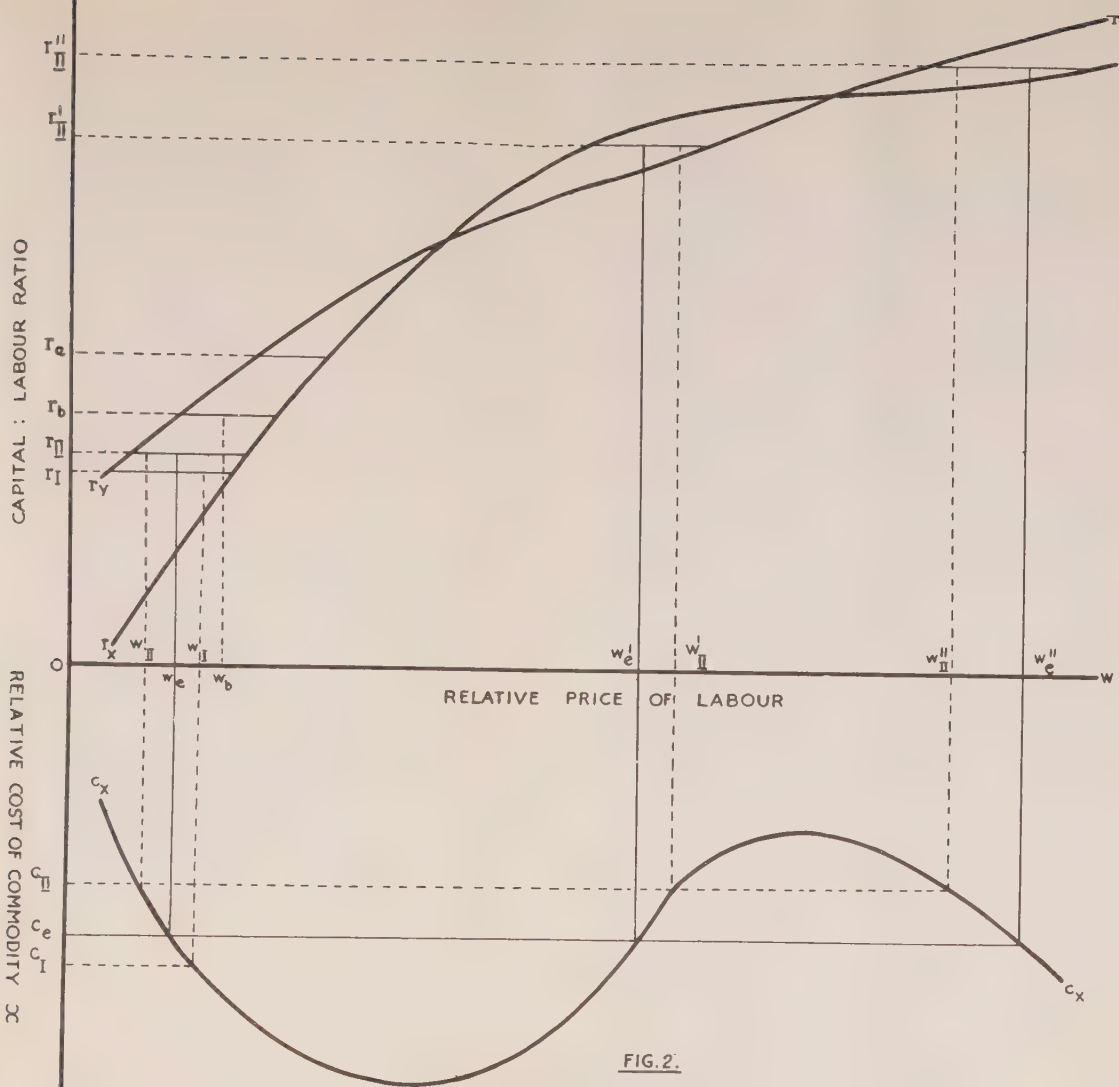

CAPITAL : LABOUR RATIO
RELATIVE COST OF COMMODITY X
RELATIVE PRICE OF LABOUR
O
w
r''_{II}
r'_{II}
r_e
r_b
r_{II}
r_I
r_y
r_x
w'_{II}
w'_I
w_e
w_b
w'_e
w'_{II}
w''_{II}
w''_e
c_x
c_{II}
c_e
c_I

FIG. 2.

ratios used in their production and the relative factor price. Little can be said about the effect of allowing for a number of countries, since this depends on the dispersion of their endowment ratios. The effect of allowing for a number of commodities, however, is to strengthen the proposition that a country will tend to export those goods which use relatively most of the factor which would be relatively cheaper in the absence of trade, and that trade will tend to equalize factor prices. The reason for this is that the more goods there are, the more likely it will be that two can be found whose relative factor intensities do not reverse at a capital:labour ratio between the endowment ratios of the two countries.

The relatively labour-intensive commodity of such a pair must be cheaper in the country with the relatively lower price of labour, and conversely, in the pre-trade situation ; and the effect of trade in establishing a common price-ratio between the pair must therefore be to raise the price of labour in the labour-abundant country and reduce it in the capital-abundant country. If both goods continue to be produced in both countries, factor prices must be equalized completely.

While the technique can be easily adapted to handle more countries and more commodities, it cannot easily be extended to allow for a larger number of factors of production, since each additional factor introduces an additional factor price and endowment ratio. Two points can, however, be made with respect to the conditions under which countries will necessarily export the commodities intensive in their relatively cheap factors, and trade will tend to equalize factor prices, The first is that, unless the number of goods is at least equal to the number of factors, no relationship between commodity prices and factor prices such as that assumed in Fig. 1 will exist, so that it will not be possible to establish conditions on which factor prices are in a one-to-one relationship with commodity prices. The second is that, once the number of factors (and goods) exceeds two, the conditions required for the one-to-one relationship to hold cannot be stated in any economically meaningful way, because the concept of "factor-intensity" loses its simplicity.

HARRY G. JOHNSON

University of Manchester.

The Legal Minimum Wage, Employers and Trade Unions: an Experiment

Contemporary wage-policy discussions have been preoccupied with two issues : the problem of restraining full employment's accompanying and persistent inflation ; and the need to introduce a greater element of rationality (variously defined) into the wage-structure as a whole. The question of *minimum* wages, which attracted many enquiries in the decades of mass unemployment and weaker trade union organisation before the Second World War, has received little attention.[1] There is a recent suggestion that some form of National Minimum Wage be adopted, but as incidental to and likely to encourage a proposed retreat from "industry-wide" negotiation to workplace bargaining in more-organised industries.[2] Generally, however, the assumption of current discussion seems to be that low wages are no longer a problem. As a result, a major development in wage-fixing—which may nevertheless have an important bearing on both the contemporary issues—has been largely neglected. Despite both high employment and trade union strength, legal minimum-wage systems have greatly extended : indeed, the curiosity is that it is just these circumstances with which the British extension, at least, seems associated. In the inter-war period, by contrast, the coverage and elaboration of legal wage-fixing each declined : the number of statutory Trade Boards, for instance, fell from 63 to 50.

The concept of the Minimum Wage appeared (when people talked about it) to imply two things : first a wage-rate or system of wage-rates separately determined and standing apart from the general structure of labour market prices, but

[1]For this reason, the standard British references are now rather antique : A. Sells' *British Wage Boards* (1939) is the most recent general study. And there is not much between that and Tawney's *Studies in the Minimum Wage*, No. 1 (Chain Making) and No. 2 (Tailoring) of 1913 and 1915 respectively.

[2]B. C. Roberts, "Is Nation-Wide Bargaining desirable," in a symposium, *The Listener*, April, 1957.

nevertheless serving as a lower limit to the latter : second, an element of state participation in the Minimum's fixing and enforcement. It would thus exclude the mere endowment of normal collective agreements or arbitration awards with some legal authority (as under the present British Industrial Disputes Order) ; on the other hand, it would equally exclude the minimum incorporated, as the lowest permissible wage-rate, in a "National Wage-Structure" (like that of Holland or, for practical purposes, Australia) the intent or effect of which is to determine the whole system of wage-levels and relationships.

Even so, the concept still permits of two forms. One is the "National Minimum" in its strict sense—the single and promulgated rate of wages below which no-one must be paid ; a close approximation to this is the United States' federal minimum wage introduced by Roosevelt. Such a minimum, however, involves a degree of state responsibility for wage-fixing that governments have often found uncomfortable : they have more usually preferred to limit this kind of socio-economic decision's political consequences by delegating it to some detached institution. Moreover, where collective bargaining is little developed, the method may inhibit its growth and encourage an embarrassing dependence on state action. And finally, a common minimum (or a schedule only crudely differentiated by, for instance, sex, age and region) applying to a great variety of occupations and circumstances, is likely in regular practice to be very low, of small effect except as a "fallback" in extremity, and to achieve little more than would, say, a comprehensive system of social insurance. The alternative assumes that low wages are normally due to a deficiency of collective organisation. It is to erect, for each trade in which it seems thus indicated, a synthetic similitude of the bargaining, conciliatory and arbitration procedures of more organised industries ; to specify what is expected of these artificial bodies or what powers they possess ; and to give their decisions a legal enforcement.

Of this second method, a very early type was the British Trade Boards, composed of people formally appointed by the Government to act the role of workers' and employers' representatives, with a number of "independent" members to

ensure that decisions were by some means arrived at. Rather similar institutions have been very widely adopted : British Colonies or former Imperial territories emulated the Trade Board, in form at least, pretty closely : but the method has been taken up quite generally by "under-developed" countries as offering both a means of wage-regulation and a half-way house to a self-supporting system of industrial relations.[1]

* * *

The British system's development, however, has itself had a changing justification. The orginal Trade Boards Act of 1909 was designed to attack notorious instances of "sweating" and could be applied to trades where the rate of wages prevailing in any branch was "exceptionally low in relation to other employments." Legal minimum wage-fixing was extended to agriculture in 1917 in essence on the principle that employers who were in receipt of public support, contracts or protection should pay "fair wages" to their workers ; it was given a limited continuance there, when the war-time subsidy was withdrawn and voluntary collective contracts failed to develop, to ensure that a farm worker's earnings should be "adequate to promote efficiency, and to enable a man in an ordinary case to maintain himself and his family in accordance with such standard of comfort as may be reasonable in relation to the nature of his occupation." But the great extension of the Trade Board system to minor manufacturing industries under the 1918 Act was designed to provide collective wage-fixing in "any trade as to which the Minister is of the opinion that no adequate machinery exists for the effective regulation of wages *throughout*[2] the trade." Though the necessity for such intervention was also to be determined ". . . having regard to the prevailing level of wages," the actual rate that might require it was left undefined ; as envisaged by the Whitley Committee (whose recommendation, of course, the 1918 Act fulfilled) its purpose was rather to promote the organisation of unorganised sectors, and as such to be an interim form only.

[1]See I.L.O. *Minimum Wages in Latin America* (1954) and *Problems of Wage Policy in Asian Countries* (1956) for many recent examples.

[2]Authors' italics.

The expectation that the apparatus of state support would wither away was unrealised. And in the latter's last enlargement, any lingering reference to the proscription of abnormally low wages was denied : the Wages Councils Act of 1945 was [1] "a declaration by Parliament that the conception of what is known as a sweated industry is past." That point had already been made, and even more specifically, in presenting the 1943 Catering Wages Bill, which anticipated 1945's general increase in both the facility and powers of legal wage-regulation. Instead, a new argument : it "was imperative," said Mr. Bevin in presenting the 1945 Bill, "to prepare for the transition from war to peace and to preserve stability afterwards. I would emphasise the word 'stability'." And again " . . . if we prevent —and I would emphasize this point—the temptation to get quick, unmeasured and violent changes in the wage system, we shall provide a great brake . . . " In effect, the new extension was justified on the ground that it would forestall the possible disruption of organised wage-relations by an uncontrolled inflation or collapse of wages in unorganised areas of the labour market.

At any rate, in the 1940s, by legislation or consequent action, the system of tripartite legal wage-regulation was extended far beyond the area of manufacturing industry. The establishment of permanent central wage-machinery in agriculture and road haulage, and particularly the inclusion of the great service trades of distribution and catering in the statutory system, gave collective wage-fixing to practically all workers not directly sheltered by normal trade union agreements—as well as a reserve procedure to many who were. In few of the trades now covered by the legal system is there no substantial sector of regular collective bargaining ; indeed, in several cases the system largely serves to impose the latter's result on an unorganised part of the industry. The number of employees within the scope of the legal Boards' orders is well over four million—about a third of the wage-earning population. And though there are nearly seventy separate statutory bodies (not including the county committees of the agricultural wage-system,

[1]Bevin, Second Reading of the Wages Councils Bill, 1945. *Hansard*, Vol. 407, pp. 69-77.

which have only secondary powers), they are clearly subject to a powerful mutual influence. Clusters of them are linked by their "representative" members being appointed from the same employers' associations or—particularly—trade unions. Of the 800 or so workers' representatives on Wages Councils 74% are currently officials or members of four unions only—the Tailors' and Garment Workers' Union, the Union of Shop, Distributive and Allied Workers, the Transport and General Workers' Union, and the National Union of General and Municipal Workers.[1] The two general unions alone hold nearly a third of all the places; these are dispersed over a wide range of boards, but the supervision of their representatives' policy is entrusted to two or three of their national officers. Similarly,[2] the 189 places for "independent members were in 1956 held by only 58 people, while of the 64 chairmanships (the most important offices) over half were filled by a bare half-dozen individuals who had charge of almost all Councils and Boards where the independent members play a significant role in the determination of wage proposals." The Boards are also linked by official and unofficial information services, and by unofficial consultation. The T.U.C. itself has its own Wages Councils Advisory Council, services the "workers' sides" of several Councils directly, and circulates a monthly analysis of statutory wage-rates from which much argument for wage-increases is drawn by workers' representatives.

The result has been virtually to establish a National Minimum Wage without calling it that—indeed, despite the T.U.C.'s repeated declaration that it was opposed to such an objective.[3] In June of 1957, the mean quartile deviation of the various rates for "all other workers" ordered by individual boards and councils (that is, the rate applicable to all employees within each body's scope for whom it prescribed no specific occupational or higher rate) was less than 4% for men

[1]Estimated from T.U.C. *Annual Report*, 1956, "Wage Councils."

[2]*The Independent Members of British Wages Councils and Boards*· F. Bayliss, in *The British Journal of Sociology*, March, 1957.

[3]The T.U.C.'s formal objects include (Rule 2): "A legal minimum wage for each industry or occupation," but as late as 1953 the Annual Congress rejected, on the General Council's advice, a motion for a Legalised Minimum Living Wage.

and about 5% for women. And the spread of statutory rates has progressively narrowed over the whole period of rising wages from 1938—at least until very recently.[1] Nor is the minimum particularly low. Even in the 1920's, trade unions were reporting that they sometimes found Trade Board Rates a useful reference in ordinary collective bargaining[2]: it is difficult to be precise but at the present time statutory wage-rates in general seem to run at between 90 and 95 per cent. of those ruling under collective agreements.[3] The legal wage-orders clearly provide much more than a "fall-back" to bargained wage-rates : they are, in fact, even more than a "floor." Statutory bodies can and do erect for their own trades quite elaborate structures of internal differentials above their lowest general rate; while their power to prescribe standard working weeks, paid holidays, guaranteed wages and so on is nearly comparable with that of a voluntary bargaining council.[4] And since the diversity of trades they cover would permit the selection of comparisons with very many well-organised occupations that would probably be accepted as valid by, say, an arbitration tribunal considering a reference from one of the latter, since employers may now be compelled to submit to such an arbitration, and since its award also has a certain legal support, the orders of statutory wage-bodies come very close to supplying a minimal *structure* of rates and conditions for wage-earners at large.

* * *

The extreme venture of the statutory wage-system's enlargement was undoubtedly its application to the catering industry. On the employers' side, there was the extent and diversity of the separate trades it embraces—pubs and clubs, hotels and boarding houses, restaurants, cafés and canteens ; there was the small scale of many businesses, so that nearly

[1]See Table I of *"Wages : . . . and the Wage-Drift,"* H. A. Turner, *Manchester School*, May, 1956.

[2]"Inquiry into the operations of the Trade Boards Acts," T.U.C. *Annual Report*, 1931.

[3]The difference is least for unskilled workers, most for skilled.

[4]The award of paid holidays by statutory bodies was, of course, specifically encouraged by the Holidays with Pay Act of 1938.

140,000 separate establishments were found to be embraced in the industry's legal definition ; and there was the relative ease of entry into several branches, combined with the frequency of working proprietorship and family employment. On the workers' side, there was the large proportion of women—over two-thirds of all employed, well over half of them married —and of part-time workers, who made up a fifth of the industry's labour[1] ; there was the high mobility of many workers, and the instability or seasonality of much of the employment—all these things reflected in a rate of labour turnover about twice the industrial average[2] ; there was the substantial proportion of workers of foreign origin ; and—particularly difficult—the element of self-employment implied in the tipping system. And as between the two sides, there was the lack of previous experience, either of such more limited wage-regulation as had held in agriculture and road haulage, or of any significantly-extended voluntary bargaining such as the co-operative societies provided in distribution.

However, it was not merely the intrinsic difficulties of their situation that made the inclusion of some three-quarters of a million catering workers the most ambitious essay in legal wage-fixing. These difficulties were to some extent recognised by giving the system a special form : the Catering Wages Act of 1943 set up a permanent Commission, from which people connected with the industry were excluded, to recommend on and review the provision of legal wage-fixing procedures for catering workers ; and the Commission's scope was extended beyond the latters' "remuneration and conditions of employment" to "provision for . . . their health and welfare and the general improvement and development of the industries in which they are employed"—thus anticipating some of the functions of the Development Councils launched by the post-war Labour Government. But the main attempt had a twenty years' history of continuing controversy : both the brief Labour minority governments of the inter-war period had moved to apply the old Trade Boards Act to catering and had fallen

[1]Approximations from the Census of Distribution 1950, and the sample returns of the Census of Population 1951.

[2]*Ministry of Labour Gazette*, monthly analysis of labour turnover.

before they could complete their efforts—the second of which was stalled by a determined litigative resistance by employing interests.

The general extensions of statutory wage-regulation have been clearly connected—whatever their immediate justification —with upsurges of trade union power and of labour's political influences. In several instances, however, they have been supported by groups of organised employers who were interested in forestalling "unfair competition" from firms not party to collective agreements : the establishment of the distributive wage-board network, for instance, was initiated by requests from the voluntary Joint Industrial Councils already operating in the retail trades. But in catering, organised employers, as well as unorganised, were predominantly hostile. Their opposition, in fact, stimulated one of the few substantial revolts against the war-time Coalition government, when 116 M.P's voted to reject the Catering Wages Bill. The Act was, in a real degree, a price of Labour's war-time co-operation. And the sense of its representing a victory in a continuing conflict to some extent persisted. "The general attitude of the trade union representatives is that they gained an advantage when the Act was passed which they were not prepared to surrender in any circumstances," alleged an employers' spokesman[1] in 1954. On the other hand, the Conservative government of 1951–55 refused to remove certain difficulties the Catering Wages Commission had encountered by amending legislation. The catering trades have thus also been the most disputed area of legal wage-fixing's advance.

* * *

The Catering Wages Act (unlike the later Wages Councils Act, which still required the Minister of Labour, at least, to have "regard to the remuneration of the workers" and to the likelihood "that a *reasonable*[2] standard of remuneration will not be maintained" in moving to apply legal regulation) quite abandoned any guiding reference to wage-policy. The Commission was simply empowered to propose the setting-up of

[1]G. Clancy, quoted in the *Manchester Guardian*, 26th March, 1954.

[2]Authors' emphasis.

Wage-boards where it considered "that machinery for regulating remuneration and conditions of employment either does not exist or is not, and cannot by any improvement which it is practicable to secure be made, adequate . . . " And five such boards were in this manner established—with a convenient vulgarisation, as follows :—

Industrial and Staff Canteen Undertakings (Canteens)
Licensed Non-Residential Establishments (Pubs and Clubs)
Unlicensed Places of Refreshment (Cafés)
Unlicensed Residential Establishments (Boarding Houses)
Licensed Residential Establishments and Licensed Restaurants (Hotels and Restaurants).

The absence of (indeed, retreat from) definitive guidance in British minimum-wage legislation induced Barbara Wootton's judgment[1] that "Statutory Wages Regulation operates, therefore, more or less in a wages vacuum, or perhaps one should say in a haze of vague suggestions, never explicitly formulated or reconciled with one another." However, examination of the Boards' actual decisions suggests they followed principles which, if implicit, are nevertheless fairly clear. Their first task was to draw up minimum wage-schedules. The earliest Trade Boards, operating in the pre-World War I context of limited and fragmentary collective bargaining, appear to have proceeded largely by reference to the current rates paid by a presumed majority of "better employers." The spread of collective organisation offered later statutory bodies a firmer reference-point, and they often ordered basic rates little below those under related collective agreements. In only one or two sections of the catering trades, however, were there such precedents as the latter ; and although in some sections again (as in the popular chain-restaurants and the brewery-owned inns), there were large employers with regular wage-scales, there was reason to think (though little statistical information on the point) that otherwise the actual wages of catering workers varied widely. In any case, the "National Minimum" already established offered a ready reference : the new Catering Boards in fact adopted (see Table I) bottom rates for men in

[1]In *Social Foundations of Wages Policy*, London, 1955, p. 86.

close correspondence to those of existing statutory bodies. The first minima they ordered for women were in fact substantially higher than those previously prevailing under legal wage-orders (see Table II). But here the Boards seem to have been influenced by the relatively high rate fixed for canteen-workers by the first Board to come into operation, and this rate in turn seems to have been based on that payable under certain voluntary agreements already existing. However, it will be noted that in the following few years the margin of catering women's rates over other statutory minima was progressively reduced—becoming negligible by 1954.

TABLE I

MEN'S MINIMUM WEEKLY WAGE-RATES

	Lowest Catering Rates											
	1. Canteens		2. Cafés		3. Pubs		4. Hotels		5. Median of all Statutory Minima		6. Agreed Rate Engineering Labourer	
	s.	d.	s.	d.	s.	d.	s.	d.	s.	d.	s.	d.
1946 ...	78	0	—		—		—		76	8	86	0
1947 ...	82	0	85	0	85	0	—		82	5	87	0
1948 ...	82	0	85	0	85	0	90	0	86	0	87	0
1949 ...	82	0	92	6	88	0	90	0	89	10	92	0
1950 ...	93	0	92	6	88	0	90	0	(92	0)	92	0
1951 ...	100	0	101	0	96	0	90	0	(98	0)	100	0
1952 ...	110	6	109	0	96	0	100	0	103	6	110	0
1953 ...	110	6	109	0	104	0	110	0	111	0	118	4
1954 ...	114	6	114	0	114	0	110	0	116	6	124	10
1955 ...	124	0	117	0	123	0	119	0	121	6	132	10
1956 ...	131	0	124	0	133	0	126	0	131	0	142	4

Source: Cols. 1 to 4 and 6 : *Time Rates of Wages and Hours of Labour* ; Col. 5 : T.U.C. (except 1950/51, which are estimated).

TABLE II

MINIMUM WEEKLY WAGE-RATES ORDERED FOR WOMEN BY STATUTORY BODIES

	Lowest Catering Rates									
	1. Canteens		2. Cafés		3. Pubs		4. Hotels		5. Median of all Statutory Minima	
	s.	d.	s.	d.	s.	d.	s.	d.	s.	d.
1946 ...	58	0	—		—		—		49	0
1947 ...	61	0	65	0	60	0	—		55	0
1948 ...	61	0	65	0	60	0	67	6	57	2¼
1949 ...	61	0	70	0	64	0	67	6	62	0
1950 ...	69	6	70	0	64	0	67	6	(64	0)
1951 ...	74	0	75	0	70	0	67	6	(68	0)
1952 ...	81	6	83	0	70	0	75	0	73	0
1953 ...	81	6	83	0	76	0	83	0	77	6
1954 ...	84	6	83	0	76	0	83	0	82	6
1955 ...	92	6	88	0	84	0	88	0	86	3
1956 ...	97	6	93	6	92	0	94	0	92	9¾

Source: Cols. 1 to 4 : Time Rates of Wages and Hours of Labour ; Col. 5 : T.U.C. (except 1950/51, which are estimated)

The Boards similarly adopted (as much an innovation in catering as a standard wage-rate) the standard working week of 47 or 48 hours then still prevalent in the legally-regulated trades. On the bases thus derived, the Boards then erected a structure of grade and area differentials. These were quite extensive : the Pubs, Café and Canteen Boards included rates for managers and manageresses, while the Hotel and Restaurant Board has issued no less than 60 separate rates for adult men and 35 for women. Such rates, however, seem to have been established largely by rule-of-thumb methods : in this case few relevant comparisons from other industries were available,

and the Boards drew largely on the job-descriptions used and differentials paid by such bigger employers as had standard scales, proceeding otherwise by consensus of opinion (implying a degree of negotiation) as to the margin appropriate in each case. The superimposing of these differentials on uniform bottom rates of course implied a considerable standardisation : but this only altered the pre-existing wage structure in so far as it reduced the latter's substantial random element. In only one important case—that of "tippable" workers—did the Boards attempt a fundamental departure. But this was a major one : they decided, on the Catering Wages Commission's recommendation, to ignore tips in computing wages, and drew up wage-scales for these workers on the basis of comparison with "non-service" grades of similar status.

The subsequent course of the Boards' wage-policy can be mainly traced from Table I. To adult men they have generally awarded increases similar to those of other statutory bodies. Up to about 1950 these increases were similar in cash terms to the equal cash increases then usual in voluntarily-organised trades, implying a persistent relative "levelling-up" of the wage-rates of workers under legal regulation. After 1950, when the more-organised industries switched to standard percentage increases, most statutory wage-fixers continued to give standard cash increases to *all* their men workers similar to those secured by the lowest-paid organised occupations, so that as a whole, their constituents fell behind—the relative recession being most marked for the higher-rated workers among them. In the last year or two (since about early 1956) some non-catering Councils with higher minima have broken away from this pattern, presumably under pressure of the widening gap between their rates and those under collective agreements, and have made quite a drive to catch up. In catering, however, the general recession was instead somewhat aggravated by an additional tendency for the Boards' adjustments to lag behind those of other wage-fixers : and this is not unconnected with the general difficulties that certain Boards have experienced in reaching agreed decisions.

* * *

The Boards thus established their initial standards by that process of nearest relevant comparison which is the essence of the "Fair Wages" principle, and modified them mainly by same rule. The greater relative recession of catering employees' rates since 1950 may, however, be associated with the particularly weak bargaining strength of the Boards' "workers' sides." And this may be reflected also in a more definite departure from general trends. In other industries—whether under voluntary or legal wage-regulation—the levelling-up of women's bottom wage-rates in relation to those of men has continued (if at a slower pace than previously) since 1950: the general tendency is towards a woman's basic wage about four-fifths that of the man's.[1] The lowest rates laid down for women under the initial catering wages orders approximated to 75% of the comparable rates for men—which was about the then general relationship (for wage-earners) outside the industry. But the Board's policy of holding back increases in women's minima until the latter corresponded to those of other statutory bodies meant that for several years the sex differential actually widened in catering. And women's wage-rates have still not recovered to their starting relationship with men's: their total increase in the whole period covered by Table II is only 72% of that awarded to adult males.

The industry's large female employment makes this a fact of some importance. On the other hand, a more obvious departure from the Boards' initial policy may be to some extent influenced by that swing of public and political opinion which may also have influenced the differing relative movements of legal wage-rates generally (as being broadly those of lower-paid workers, as against the better-paid of more-unionised industries) before and after 1950. The role of the statutory bodies' independent members, if not quite that of compulsory arbitrators, certainly goes beyond that of mere conciliatory mediation, at least on boards where the collective organisation behind the "representative" members is particularly weak: and there are several indications that the independent members provide an avenue through which outside opinion may influence

[1]See *Manchester School*, May, 1956 (*op. cit.*).

statutory decisions. At any rate the original precept that tips should be ignored in fixing wage-minima, adopted in 1947 by the Catering Wages Commission after an inquiry to which it was directed by the Ministry of Labour, implied a minor social revolution. Recognition of tips as part of wages, however correct in an economic view, involves of course formidable difficulties for systematic wage-regulation of any kind : but the driving argument—expressed clearly by the Labour majority in Parliamentary discussion—was that the tipping system was degrading to human dignity. And the hope was that its non-recognition would induce it to wither away.

However, the tipping system's recalcitrant survival upset the new wage-structures' balance as between tippable and non-tippable workers. It soon became clear that its elimination would require a more positive attack. The employers were pressing for action, but the unions had refused to join them in a public campaign against the system : in 1950, however, when the Commission again considered the problem, it rejected their proposals of various statutory remedies, and returned instead a recommendation (that they and the unions negotiate an agreed service charge, to replace voluntary tips and from which wages might be paid) which was under the circumstances doomed to ineffectiveness. The drive behind the original decision was apparently fading. In 1952/3 certain Boards awarded wage-increases to non-tippable staffs only, so that tippable workers' rates were relatively reduced. The device compelled a more permanent (if still unsatisfactory) solution. And in 1954 the Hotel and Restaurant Board authorised deductions (of up to 20*s*. a week in the first place) from the minimum wage-rates of tippable workers, subject to the latter's individual consent. Tipping, in effect, was now accepted.

* * *

The Catering Boards' wage-policy has thus been arguably influenced by several factors : the comparative concept of "Fair Wages" which is generally the most important social pressure in wage-determination ; the differentials established by a relatively free labour market before legal wage-fixing's introduction ; the degree of bargaining organisation of workers

and employers ; and the general climate of opinion on social/economic matters. But just how much effect did Board decisions have on wages actually paid ?

General impressions are that the original Board rates were below those often paid in canteens, hotels and restaurants, and about the predominant standard in pubs, clubs and cafés. There is no doubt, however, that the Boards' minima have exercised a considerable "levelling-up" influence. The clearest evidence for this is the unusually high degree of underpayment on their legal rates which the Boards' inspectors continue to discover and reclaim.[1] Table III shows that of the establishments examined (in the first years of legal regulation, about a fifth of those to which it related, but declining to about a tenth in each later year) a fairly steady quarter have been found liable for underpayment, and that this affected about 10% of all the individual pay-records examined.[2] These figures are, of course, liable to considerable qualifications because they include large elements of both under-counting and over-counting ; on the one hand, the many cases which the inspectors did not pursue because of their triviality, unsatisfactory evidence, or the workers concerned having left (a particular likelihood in catering) are not reported ; on the other hand, underpayment under separate heads of basic wage-rates, overtime (including "spreadover") rates, and holiday pay is separately counted, so that the same employer or worker may be included twice or thrice. Moreover the sample is not random, because a large proportion of inspections are made on "complaints" and not on "test." In fact, however, considerable variations in the ratio of "complaint" to "test" visits have apparently had little effect on the degree of underpayment reported. While most of the underpayment—since a normal working week is rather abnormal in catering—is heavily concentrated on overtime/"spreadover" rates. This, of course, raises the question of how far the complexity imposed on wage-orders by the catering trades' working conditions and the requirements of legal precision leads to accidental underpayment: and it is certainly the case (as the figures suggest) that

[1]For comparisons with other statutory wage sectors, see *Manchester School*, May, 1956 (*op. cit.*).

[2]The 1946/7 data is mainly for canteens only.

TABLE III

ADMINISTRATION OF THE CATERING WAGES ACT, 1943. 1946–1955

	1946	1947	1948	1949	1950	1951	1952	1953	1954	1955
No. of Establishments on lists, end of year	24,379	123,476	132,959	134,763	138,025	139,546	139,566	138,766	136,404	134,695
No. of complaints	174	574	2,144	3,057	3,138	3,072	3,318	4,400	3,900	3,951
No. of Establishments inspected	5,538	22,899	25,969	22,312	11,663	11,972	13,062	13,302	14,301	10,584
% of Establishments by which arrears paid ...	12	15	23·3	25·6	25·4	23·2	25·6	25·2	23·2	24·2
No. of workers whose wages were examined ...	25,008	94,468	145,398	121,134	63,213	59,909	68,620	70,667	75,897	57,708
% of workers to whom arrears paid ...	8·5	8·1	11·7	11·7	10·7	9·6	10·8	10·2	9·3	9·2
Average amount of arrears recovered (£) ...	2·4	3·2	4·4	7·5	9·3	7·8	6·9	6·6	7·1	7·3

Source : Annual Reports of the Ministry of Labour.

under-payment is largely attributable to smaller establishments which could not usually have a specialist in such matters. But on general experience, smaller firms are the worse-payers anyway. Many smaller catering establishments formerly had to pay higher wages because tips were usually lower in such places: and it used to be customary in catering to pay a flat rate to cover the whole working time. Where this practice has continued, and the flat rate covers the total sum the relevant legal rates would require over a reasonable period, it has been accepted by the inspectors. Both standard and supplementary rates in catering are low by outside comparison : and one would expect the abnormal working hours to be reflected in higher pay. There seems no reason, therefore, to doubt that those catering establishments found guilty of under-payment are in fact paying low wages, and that the general "levelling-up" is real enough.

On the other hand, there seems also to have been a certain levelling-*down* to Board rates. There is a lack of catering earnings data for comparison, but an estimate for 1950 can be made from the Census of Distribution. Weekly earnings (excluding tips but including meals and lodgings—for which allowable deductions from Board rates are specified—and other bonuses) then averaged 89*s*. 6*d*. in cafés, against a median of all the relevant statutory rates *for standard hours* of 82*s*. In canteens they averaged 83*s*. 4*d*. against a relevant median[1] of 89*s*. Despite the comparisons' approximate nature, they stand sharply against a general labour situation in which average earnings appeared to be about 20% (roughly half of which margin was accounted by overtime) above standard wage-rates.[2] The levelling-down appears to have taken two forms. First—and here one must rely on scattered reports and the impressions of officials and others—firms which were at first paying above the initial Board rates often held their own wages steady thereafter, until the former rose to the latter. A tendency of employers to treat statutory minima also as conventional maxima has been noted in other trades where unions are

[1]It seems impossible to compute a weighted average of rates. Part-time workers counted as halves in computing earnings.

[2]See *Manchester School*, May, 1956 (*op. cit.*).

particularly weak—especially in agriculture (in manufacturing the picture is confused by the prevalence of piecework).

The other form of levelling-down is more interesting : in several cases groups of catering workers are attached as auxiliary to some other, and voluntarily-organised industry. An obvious case is industrial canteens : here two unions had an agreement before the 1943 Act with the association of contracting caterers, but these employed only a small minority of the 200,000 workers. In engineering establishments it was otherwise customary to relate canteen staffs' wage-rates to those of unskilled production operatives. When the Canteen Board began to operate canteen workers were excluded from engineering wage-increases. Many firms then either kept their canteen wage-rates stable till legal wage-rates overhauled them, or applied only the subsequent legal *increases* to their existing canteen rates, so that the latters' differential over legal minima was maintained but they fell behind production operatives' wage-rates. The last technique was quite widely adopted by public authorities in other cases : the wages of employees in the Carlisle state pubs scheme used to be related to those of Post Office workers, but when the Pubs Board was set up they received only the increases it ordered. Adjustments in the wages of employees in Civic Restaurants, once governed by the local authority manual workers' pay, were similarly tied to the increases in Café Board minima. Even the wage-rates of some workers outside the legal Boards' scope, like those in Ministry of Supply canteens, had their connection with voluntary agreements severed and re-tied to statutory wage-movements. In all these cases, employers had formerly applied the "Fair Wage" principle to their catering workers. They continued to do so : but the establishment of legal minima offered them an alternative, and cheaper, comparative reference. And arbitration tribunals apparently found it a valid one, since unions appealed unsuccessfully to the Industrial Court or Disputes Tribunal in several such instances.

How far the various tendencies for pre-existing higher rates to be levelled down to statutory standards offset the latters' levelling-up effects cannot be reckoned. And there was, of course, the unforeseen impact of the first decision to ignore

tips. Though the majority of catering workers do not receive them, there is no doubt that the actual income of many who do was for a time substantially boosted—particularly among the higher-earners in luxury establishments. There is an impression that tips are relatively of less importance, which if correct suggests that public knowledge of the minima's existence may be slowly undermining the tipping system just when the Boards have come to recognise it. In that case the ultimate result should prove interesting; however, it seems quite impossible to estimate the present effects of tipping on earnings.

* * *

The broad effect of the Catering Boards' activity was to level-up—and to some extent level-out—wage-rates to standards comparable with those in other legally-regulated occupations. What impact did this have on the industry itself? We have no information which would indicate any changes in the total employment of the catering trades associated with the Boards' establishment, since that coincided with general alterations in the coverage of public employment statistics. There are suggestions in subsequent series of a reduction both in seasonal employment fluctuations and in labour turnover: however, the last is a fairly widespread phenomenon and may be attributable as much to a general settling-down of the labour force after its immediately post-war disturbance as to specific changes in conditions. But it is notable that the number of catering establishments in the Boards' scope shows a persistent decline from 1951 onwards (the previous *increase*, of course, largely reflects the Inspectorate's completion of its lists). There is some direct evidence that legal wage-regulation stimulated efficiency: many smaller establishments, in particular, were not very adept at deploying their labour in the most economical way, and the Boards' Inspectorate has regularly given advice on such matters as the arrangement of staff working hours to minimise "spreadover" rates, etc. But the more important effect is indirect—the pressure of the levelling-up in wage costs on less efficient establishments.

The appearance of sharp competition in catering, both for trade and labour, is to some extent misleading. The trade is certainly competitive in, for instance, the centres of big towns : but otherwise many small establishments have minor local monopolies ; and in such places as holiday resorts, which are subject to intense short period demand, it does not seem unusual for common price-scales to be applied by arrangement or understanding. On the other hand, much of the apparent labour mobility is internal to the trade—and a good deal of the labour is not mobile at all. Under such circumstances it is very possible for an inefficient use of labour to develop simply because it is too cheap. It can well be argued that the Boards, by supplying a substitute for the pressure exerted by trade unions in other industries, have stimulated efficiency in the catering trades. That some such pressure should be put upon the service industries appears trebly important : they account for an increasing share of total national output and capital ; their natural rate of increase in productivity is negligible ; and they generally employ the least-organised workers.

* * *

The question, in fact, is whether the Boards exerted pressure enough ? The tendency for statutory wage-rates in general and those of the Catering Boards in particular to lag behind rates in more organised occupations after 1950 implies that the pressure was relaxing. And the lag was most marked for the rates of more skilled and experienced workers, who—being also more committed to their trade than those in lower grades—are naturally less mobile.

These things we have attributed to the bargaining weakness of the workers' sides, an interpretation which seems confirmed by the tendency of certain statutory bodies—and particularly those in the distributive trades, which have been quite aggressively used by U.S.D.A.W.—to pull away from the generality of legal awards in the last year or two. It is not necessarily the whole story : for instance, the relatively greater recession of skilled catering employees' rates after 1950 may be in part attributable to the unions' persistence, until 1953/4, in flat-rate demands—a policy characteristic of the general unions.

These demands were apparently satisfied with increases comparable to those secured by the lower-paid in other industries, which were then applied to *higher*-rated catering workers. But to a weakness of collective organisation certain more definite failures of the Catering Wages system can be traced. There is, of course, the failure to solve the tipping problem. This was central to the whole question of wage-regulation in certain branches : and its solution would have been one of the most signal public services that the Catering Boards could have performed. The Commission's proposal for an agreed service charge, however, implied a degree of collective organisation and discipline far beyond that actually existent.

Particularly, there is the difficulty experienced by the two Boards originally established for Hotels and Restaurants, and for Boarding Houses. These two Boards—unlike those for Canteens, Cafés and Pubs—had virtually no previous background of limited collective bargaining. And their negotiations were clearly hampered by the unrepresentative quality of the "representative" members. This seems to have been particularly the case on the employers' side, where it was not merely that very many caterers belonged to no association, but that the associations tended to be dominated by the larger firms. For instance, the pleas of smaller proprietors that a single, just and simple form of wage regulation could not be devised to suit the luxury hotel, the small hotel and the seasonal establishment, and that a division should be made by size, were ignored. Said the Commission in its own review of the situation[1] : "Not a few of the difficulties to which the Licensed Order [2] was claimed to have given rise, resulted in our view from two causes. Firstly, upon the employers' side the needs of the smaller establishments were overshadowed by those of the larger and luxury units. Secondly, the agreement intended to prevent premature notice to the Press of matters from time to time under consideration by the Boards was so construed that proprietors of the smaller concerns were never consulted as to the course negotiations were taking, negotiations which would affect their

[1]Report of an inquiry into the operations of the Catering Wages Act, 1943, in the hotel industry, by the Catering Wages Commission. Cmd. 8004. H.M.S.O. 1950, pp. 20-21.

[2]*i.e.*, that applying to Hotels and Restaurants.

interests as vitally as those of the larger units. Indeed, some who came before us, and who had taken exception to the Licensed Proposals when published, stated that they were led to believe that such Proposals had originated with the Independent Members who, upon all matters of principle, had on every occasion voted with workers' side and that a refusal to accept them could in the circumstances only result in proposals more stringent. That such a belief was unfounded is we think shown by the record of the Board, and we regret that the representatives of the employers' side did not keep in touch with the smaller units in the industry."

On the Hotel and Restaurant Board, in fact, many meetings broke down because the employers' representatives could not agree among themselves—for instance, on whether a telephone operator was a service worker. On one occasion the employers' side withdrew from a meeting which passed a wage-claim of 10*s*. 0*d*. a week in their absence. While the record of the Boarding House Board on the specification of overtime and "spreadover" payments (Tables IVA and B) shows how the employers representatives on that body retreated from decisions to which they had once subscribed :

TABLE IV

RECORD OF DISCUSSION ON HOURS OF WORK.
UNLICENSED RESIDENTIAL WAGES BOARD.

A. **Hours and overtime.**

Date of Wages Board Meeting	
22nd January, 1947	AGREED TO BY BOTH SIDES—48 hours to be normal week.
19th May, 1947	AGREED TO BY BOTH SIDES—Overtime in excess of 48 hours. First two hours—Time and a quarter. Thereafter —Time and a half.
15th July, 1947	EMPLOYERS' PROPOSAL—Overtime in excess of 48 hours. First six hours—Time and a quarter. Thereafter —Time and a half.
	WORKERS' SUGGESTION First three hours—Time and a quarter. Thereafter —Time and a half. The workers' suggestion was accepted by the employers' side and adopted by the Board.

TABLE IV—continued.

15th April, 1948	EMPLOYERS' PROPOSAL—Overtime to be calculated on a fortnightly basis. Decision deferred.
20th October, 1948	EMPLOYERS' PROPOSAL—Overtime to be payable after 96 hours and 60 hours in any one week. AGREED TO BY BOTH SIDES—Overtime in excess of 96 hours in a fortnight. First six hours—Time and a quarter. Thereafter —Time and a half. In excess of 56 hours in a week—Time and a half.
31st May, 1949	EMPLOYERS' SUGGESTION—First 12 hours in excess of 96 hours in a fortnight —Time and a quarter. Thereafter —Time and a half. No provision for weekly hours.
	B. Spreadover.
22nd January, 1947	EMPLOYERS' PROPOSAL : The daily hours of duty may be spread over a period of 12 hours in the case of non-residential workers and 15 hours in the case of residential workers, inclusive of meal times and a period of not less than three consecutive hours of duty.
19th May, 1947	EMPLOYERS' PROPOSAL : (*a*) Split duty is a turn of 12 hours or more, which includes an interval of two hours or more, apart from meal times ; (*b*) for each split duty a worker shall be paid 2d. for each turn worked, in addition to the appropriate rate of pay ; (*c*) for each split duty of 14 hours or more a worker shall be paid 3d. instead of 2d. This proposal was agreed to at meeting of theWages Board on the 15th July, 1947.
29th October, 1948	EMPLOYERS' PROPOSAL : Provisions for spreadover should not be implemented for twelve months. Defeated.
17th December, 1948	EMPLOYERS' PROPOSAL : Provisions for spreadover should not be implemented. Defeated.
31st May, 1949	EMPLOYERS' PROPOSAL : First that the spreadover provisions be omitted altogether, and then that provision for spreadover of 12 to 14 hours be deleted. This was carried with the support of the independent members.

The weakness in collective organisation, and the lack of experience in negotiation, were in fact such as to put the independent members of these two Boards in an embarrassing position. Many deadlocks occurred in negotiations, thus forcing the independent members to vote with one side or the other. Independent members cannot, of course, initiate proposals but they can influence the course of negotiations (one workers' side in 1955, for instance, withdrew a demand for a 44-hour week on the understanding that the independents would support a wage-increase). Both sides, however, apparently resented the independent members' power.

It was remarkable that in concluding its review of this situation, the Commission—having recommended a single Board for all residential establishments, with sub-committees to propose wage-regulations according to the size and situation of hotel—proposed : "That the membership of such sub-committees, if set up, should be confined to persons actually and actively engaged in the hotel industry and the members should not be drawn from the employers' or workers' representatives on the Board." [1] It has always been the practice of Wages Boards and Wages Councils to recruit nominees from trade unions and employers' associations. In the case of workers there has often been the fear of victimization : in the case of employers there has been the difficulty of obtaining people with the qualities required for a negotiator. But the main reasons for drawing recruits from organizations have been the desire to stimulate and foster normal collective bargaining proeedure, and sheer convenience of management. To staff Boards from the unorganized can only lead to too much discussion and too little action, because no one knows the rules : organization gives responsibility and stability. Viewed in this light the Commission's recommendation could have only been made on the grounds that there was no adequately representative organization to draw on.

The Commission's proposal for reconstruction encountered some legal difficulties and was revised to one for four separate wage-boards. But this proposal too met a legal obstacle. In the upshot the Commission reconstituted the Hotel and Restaurant

[1]Hotel Inquiry, *op. cit.*, pp. 40-47.

Board with several functional sub-committees ; but only one of these has at writing issued proposals that have been embodied in a wage-order. The Boarding House Board has not been reconstituted, and the obstacles to any regulation of unlicensed hotels have apparently grown. In this sector of catering, the weakness of collective organisation led not merely to a relative decline in statutory conditions, but to an absolute breakdown of the attempt at legal regulation.

That the examples here quoted largely illustrate an ineffectiveness of employers' organisation is beside the point. Since such organisation generally grows in response to that of trade unions, its weakness here can be taken to reflect the failure of trade unionism itself to develop. The employers, on the whole, did not want the Catering Wages Act to work anyway : but had they been confronted with an effective union organisation, they would soon have produced more adequate negotiating agencies.

* * *

What all this suggests is that so far from a system of legal wage-regulation serving as a substitute for voluntary collective organization, any such system that attempts to govern at all closely the conditions of particular occupations, itself requires a substantial degree of parallel (if not pre-existent) voluntary association. In the absence of that condition, only the cruder sort of National Minimum—the simple standard rate propulgated without reference to the circumstances of individual trades—appears practicable.

The thesis that a phase of legal control provides an interim form, which would inevitably encourage such voluntary organization as would make statutory regulation ultimately unnecessary, seems equally somewhat discredited by this experience. On the employers' side in catering there certainly appears, since the passing of the Catering Wages Act, to have been a growth in organization as well as a decline in the number of associations as a result of merger. In 1952 the Caterers' Association had 8,000 members representing 20,000 establishments ; the Engineering and Allied Employers' National Federation is also important in industrial catering, and the Licensed Trade Defence Association in pubs. But

outside these groups (in which legal regulation also appears to have proceeded relatively smoothly), it was fairly clear that employers' organization was still pretty unrepresentative in character. The British Hotels and Restaurant Association (B.H.R.A.) was formed by the amalgamation in 1948 of the Hotel and Restaurant Association of Great Britain and the Residential Hotels Association of Great Britain. A letter circulated to the members of both associations advocating their merger stated that "with the introduction of Government controls and the growth in the power of the trade unions, the employers must have the strongest possible organization which can speak with unmistakable authority on behalf of the catering industry." In 1953 B.H.R.A. announced that there had been an increase in membership of 183 employers, whilst the number of affiliated associations had risen to 79. By 1955 B.H.R.A. represented the interests of 10,000 establishments and 83 affiliated associations. But the great bulk of caterers remain, apparently, unassociated. The stimulus first given by the Catering Act itself faded as it became clear that no major threat of unionization was to follow. The effect of legal regulation on trade unionism seems, indeed, to have been negligible. And this is curious, since it was from the unions in the main that the pressure for statutory organization came, and one might perhaps have expected a more determined attempt to exploit the opportunity it was alleged to create.

Such catering workers as are organized (about 10 per cent. of the labour force) are members of twelve unions. However, the principal of these are the Union of Shop, Distributive and Allied Workers (U.S.D.A.W.), the National Union of General and Municipal Workers (N.U.G.M.W.), the Transport and General Workers Union (T. & G.W.U.), the National Union of Railwaymen (N.U.R.). The unions have avoided the internal competition that troubles some employees' organisations. They co-ordinate their activities through a Joint Trade Union Committee for the Catering Industry, which was established following the passing of the Catering Wages Act. One of the first tasks of this Committee was to lay down spheres of influence for its member unions. " . . . we created an understanding between 12 Trade Unions in the country covering England,

Scotland and Wales, and we have been congratulated by those whose experience of the Trade Union Movement is great, much greater than my own. The fact is that spheres of influence were bisected to prevent so many of those heartburning situations we find in other industries, where we have the spectacle of Unions fighting one another on lines of demarcation, a hell of a sight worse than they fight the employers. We have divided England, Wales and Scotland into a variety of areas. We have agreed that spheres of influence should be allocated to specific unions." [1]

The success of these jurisdictional demarcations must, however, be mainly attributed to the fact that the catering industry is still largely virgin territory. Friction only arises when unions are forced up against each other in strongly organized industries. There has, indeed, been little union campaigning. The large areas blocked out by the agreement have been regarded by the unions as valuable but merely potential collecting grounds : so far little has been done to explore them. A certain amount of recruiting was done in the immediate post-war years. The N.U.G.M.W. carried out a successful drive in the London area during 1946-7. After the Savoy strike, however, membership fell off again, partly because of the failure of the strike itself and partly because of the Executive's firm action to suppress what appeared to be a dissident, if militant, movement.[2]

Catering workers are, of course, not easy to organize. Organizing costs are likely to be high, and the general unions, which are particularly concerned, may be reluctant to spend to such poor return money contributed by their other groups of membership. Before the Catering Wages Act, the T. & G.W.U. had already asked to be released from its exclusive responsibility for organizing catering workers in London. The availability of legal regulation may well have diminished the pressure to undertake such a task—a task, incidentally, for which the modern trade union administrator and negotiator is not usually well-equipped by experience. From the general unions' viewpoint,

[1]Report of the Thirty-Second Congress of the N.U.G.M.W., 1947, Speech : T. Cochrane, p. 128.

[2]c.f., H. A. Clegg, *General Union*, Blackwell, 1954, pp. 124-30.

the catering trades could well represent an area of threateningly non-regulated conditions, a reservoir of potentially-competitive labour rather than a reserve of union strength. Legal regulation, as it were, secured their flank, and made expensive organizing expeditions into this difficult territory unnecessary.

It may thus be that the catering wage-board system appears, to the unions at present mainly concerned, to do as much as they really require of it. These unions, however, have (as their titles suggest) their principal foci of interest outside the catering trades. It could be argued that a union primarily of catering workers would have made a more forceful attempt to exploit the statutory machinery—as the U.S.D.A.W. has utilized the distributive Wages Councils, and the Tailors' & Garment Workers' Union once capitalised on the old Trade Boards. Such a union would be obliged to look first to an aggressive recruiting campaign to safeguard its own existence, and would be compelled to seek collective agreements to improve its bargaining position on the wage-boards themselves. It might well attract those more skilled catering workers who fear a loss of their trade identity in the general unions. Some might add that it would probably show a greater understanding of the specific circumstances of the catering trades than unions whose main experience lies in other industries : it has been suggested to the writers that—although these unions naturally have their specialist officials—some of the difficulties of the catering Boards have sprung from an attempt to transfer too mechanically to them the standards established in occupations of a very different character. It could also be argued that an independent catering workers' union would have displayed greater resolution in approaching certain specific problems of the catering industry—such as that of tipping—because an enforceable policy on such questions would be central to its successful operation.

Be that as it may, the idea of an independent occupational union has been raised and defeated on several occasions. As recently as 1951 (and presumably as an aftermath to the Savoy affair), the following motion was rejected by the N.U.G.M.W's 36th Congress : "That owing to the failure of the N.U.G.M.W. to organize catering workers to any substantial

degree, the National Executive Committee be instructed to conduct a ballot to all members of the catering trade as to the advisability of carrying on membership within the N.U.G.M.W. Should a two-thirds be negative, then the National Executive shall recommend to the T.U.C. that the catering workers should be allowed to carry on forthwith to organize catering trade workers into a National Union of Catering Trade Workers." And an independent catering workers' union that functioned in the 1920's—the International Geneva Association of Hotel and Restaurant Employees—has long since become extinct.

In truth it must be said that an occupational union formed now would be confronted by a double difficulty, which could only be overcome with the help of the trade union movement in general. First, it would be lacking in a core of solid organization and in financial resources. Second, it would run up against the doctrine of "organizing rights" which customarily regulates British inter-union relations, and would be unlikely to achieve recognition without a lengthy struggle. Since the unions at present concerned with catering would apparently be unwilling to surrender their organizing rights to an independent organization, its practicality seems likely to remain an academic question. The recent affair of the Northern dockers' breakaway from the T. & G.W.U. has shown that, even where a firm tradition of occupational organization already exists, the obstacles to any new union grouping are well-nigh overwhelming. But that being so, it does seem possible that more could be done by the unions now engaged : in one case encountered by the writers one of these unions had neglected to organize—even to approach—the staff of the café next door to its district office.

* * *

To summarise this survey's results. The catering wage-boards represented—with the subsequent extension of the Wages Councils system to distribution—the last bricks in a structure of legal regulation that now supports the pay and conditions of virtually all wage-earners whose basic rate is not secured by firm collective agreements. Only private domestic service remains outside the intended scope of some kind of collective regulation—voluntary or statutory—and that, as a

field of full-time employment, has dwindled to comparative insignificance, anyway.[1]

How one assesses the catering experiment naturally depends largely on the viewpoint one takes in social-economic matters generally. The introduction of legal wage-fixing procedures, however, has generally been prompted by two motives : first, to establish basic wage-standards which are comparable with those generally prevailing : second, to stimulate voluntary collective organisation of employers and workers, and to prepare the way for a self-supporting system of collective bargaining.

Under the first head, the functioning catering wage-boards have undoubtedly effected a considerable levelling-up of wage-rates ; and this has probably had the incidental effect of reducing that subsidisation of inefficient firms by cheap labour which an imperfect labour market still tolerates in the presence of generally high employment. On the other hand, the introduction of legal wage-fixing led to a significant levelling-down of wage-rates which were higher than those provided by the wage-boards' orders. Moreover, in several respects, the catering Boards have fallen back from the standards they first established, so that these are becoming less favourable, not merely by comparison with those of workers under normal collective agreements, but even in relation to those provided by other statutory wage-fixers.

These latter trends are due to the catering wage-boards' relative failure under the second head of their motivation—the encouragement of voluntary collective organisation. Under the Catering Wages Act, employers' associations have developed somewhat, trade unions only a little : neither among employers or workers, however, has organization yet grown to the point at which it can be considered even substantially representative, except in one or two branches where it was already established. And voluntary collective agreements have hardly increased at all. On the boards themselves, apparently, bargaining power is still important, and the position of the workers' sides seems

[1]Even there, of course, the National Institute of Houseworkers represented an attempt to introduce an alternative form of collective influence.

relatively weak. While in one major sector voluntary organization has proved insufficiently extensive to permit the wage-board system to function effectively at all. It has become clear, in effect, that—contrary to legislative theory—a system of legal minimum wage-regulation that aims to provide such detailed standards as does the British itself implies the existence of unions and employers' associations which are experienced in negotiation and can command the support of a substantial proportion of the firms and workers affected.

In this weakness of the catering Boards, the key factor has been the negligible growth of trade unionism. The catering trades provide a poor field for trade union organisers—which is why legal regulation was thought necessary in the first place—but the lack of any determined attempt by the unions to exploit the position they secured under the Catering Wages Act encourages a suspicion that they are not so interested in recruitment there as to undertake its trials and expense. These unions are based on other industries, to which the catering trades are merely contiguous ; it could be that the unions saw the catering boards rather as a device to protect their established interests than as a channel to extend them. If so, the calculation may yet prove mistaken.

* * *

By way of conclusion, however, we can perhaps consider the possibility that a system of legal wage-regulation itself discourages trade union recruitment. If real, this would of course deny a major justification of the minimum wage-system. Nevertheless—and curiously, because the pressure of organized labour has been the main force in statutory wage-fixing's recent extension—the proposition is quite commonly argued in trade union circles. Workers are supposed to think union membership unnecessary because the state protects them. "We have repeatedly heard," said Mr. Arthur Deakin at the 1953 T.U.C., in opposing on the General Council's behalf a motion for a single legal minimum, "that the Wages Council militates against good and effective trade union organisation."

The argument seems of doubtful validity. There are now not a few cases where an efficient voluntary negotiating system

is combined with legal regulation—in several garment and distributive trades, for instance. Even in catering, the railway hotel workers have preserved separate collective agreements which provide conditions superior to the legal minima. If legal regulation *does* handicap trade union recruitment, the unions have shown no anxiety to rid themselves of it, despite (now) nearly half-a-century's experience. In 1952 the Ministry of Labour proposed that legal regulation be suspended in selected trades for an experimental period. Over two years' discussions, the T.U.C. General Council reiterated its desire to see legal wage-regulation ultimately give way to effective voluntary negotiating machinery, but rejected the proposed experiment.

The legal wage-board in fact offers several advantages to unions operating in industries which are otherwise hard to organize. It gives them both a public recognition and a public platform, and compels employers to meet them. And the system involves no hindrance to their approaching workers which a propagandist instruction seems incapable of resolving. The unions are still required to represent workers on the wage-board itself, while a system of enforcing wage-orders under which individual establishments will only be visited by public inspectors (on the average) once in fifteen years clearly leaves much to voluntary action.

It is hard to avoid, in fact, a sense that in putting forward (publicly—their private policy is apparently more realistic) such criticisms of the legal wage-board system as the above, union leaders are perhaps constructing an alibi for themselves. The system apparently permits them to participate in fixing wages for more stubborn sectors of the labour market without taking the trouble to organize the workers concerned. Nor is there any great pressure on them to do so, so long as the boards themselves continue to apply the "Fair Wages" concept and maintain statutory rates which are nearly comparable with those in organized industries. Should, however, the need to recruit in these sectors arise in the future—or should any unforeseen and spontaneous enlistment there occur—the British trade union movement's doctrine of "organizing rights" meanwhile virtually guarantees them as permanent reserves of established unions. Till then, the wage-board system itself

offers a convenient scapegoat for the unions' present failure to recruit.

The real risk of the statutory system is thus that it may make trade unions lazy. And this possibility raises a general question. It is curious that British trade union membership, despite continued high employment and generally favourable conditions, should for the past few years have been relatively stable at about the figure it attained at the end of the immediately post-war uprush. That figure (some nine-and-a-half million) represents less than half of all employees, and indicates a substantially lower degree of union organization than in several other countries—Scandinavia, for instance : it is probable, even, that less than two-thirds of British wage-earners proper are union members. There is no evidence that union leaderships are particularly disturbed by this situation. At no point in the past decade have the major unions undertaken substantial recruiting campaigns : and the trade union movement seems persistently neglectful of its public relations in general. How far is this seeming complacency attributable to a sense on the unions' part that any threat to their established standards from unorganized areas of the labour market is warded off by the elaborate system of state support that operates in the latter ?

If that consideration is at all influential in contemporary union leadership, it clearly involves a double danger. First to the quality of the legal wage-fixing system itself. It is not merely that the most hopeful line of development from the latter lies in its exploitation by the unions to create new areas of collective bargaining : the present study suggests that *unless* this occurs, the legal system itself can be of only limited effectiveness. The second danger is to the unions' own internal health. Trade unions have been charged with bureaucratic tendencies by people who think their leaders do not consult their members closely enough. How much more substance would the charge have if union leaders need no longer bother to acquire members to consult !

University of Sheffield. BRIAN MCCORMICK

University of Manchester. H. A. TURNER

A Note on Secular Swings Around Alternative Trends

In their investigations of the interaction of the United States, European and a number of other economies between 1870 and 1913 [1] Lewis and O'Leary demonstrate the prevalence of 16–22 year swings in economic activity over this period. Their findings are based on the study of the behaviour of deviations from a straight line trend fitted to the logarithms of some 45 volume series and, in their introductory remarks, the authors note that : "Fitting a straight line is . . . . . misleading if there is constant retardation or acceleration in the rate of growth." [2] An attempt by the present writer to apply Lewis and O'Leary's method to determine whether similar swings were to be found in competing imports [3] has convinced her that the danger they mention is a serious one and that Lewis and O'Leary themselves may have taken it too lightly. In the following remarks the theoretical problem is discussed and subsequently illustrated by examples from Lewis' and O'Leary's series and from the writer's own work on competing imports.

The problem is to decide how the underlying growth tendency of a series can be approximately represented. We wish to decompose the series into a growth trend and the cyclical fluctuations of the observations around this trend. The greater the extent to which a series is trend dominated the easier it is to decide which form of time function best describes the manner in which the series is growing. Given the average duration of the particular cycle we suspect, the longer the period of time for which we have observations the better the chance any trend covering the period has to assert itself. The trend is also easier to discern when the amplitude of the cycle is constant and small relative to the average rate of growth over the entire period. The obvious difficulty with most

[1] Lewis and O'Leary, "Secular Swings in Production and Trade 1870-1913," *Manchester School*, Vol. 23, No. 2, May, 1955.

[2] *Ibid.*, p. 115.

[3] Imports of Foodstuffs competing directly with the products of British Agriculture.

time series is that they are not available for as long a period as we would like. If a Production or Trade series is subject to wide cycles of 16–22 years duration it is unlikely that observations for a few decades only will display an unambiguous growth trend. It is, therefore, in the investigation of 16–22 year cycles that the problem of choosing a trend function to fit the series is particularly difficult.

Of the various methods of representing the trend it is sufficient for the present purpose to confine our attention to the semilogarithmic polynomials in time. These have the advantage that the constants of the functions can be determined by least squares so that, provided that we can assume that growth is subject to the conditions implied by the function chosen, the method of estimating the values of the constants ensures that the trend will pass as closely through the centre of the observations as is possible for a function of that form. Furthermore the rate of growth of the series is yielded by the first derivative of the function. That is to say, if Y denotes the trend values and if $y = \text{Log}_e\ Y$, the rate of growth of Y is given by

$$\frac{1}{Y}\frac{dY}{dt} = \frac{dy}{dt}$$

Thus it is convenient to consider y as a function of t.

In empirical work we are concerned almost exclusively with polynomials of the first, second, and third degree.

(i) $y = a + b_1 t$
(ii) $y = a + b_1 t + b_2 t^2$
(iii) $y = a + b_1 t + b_2 t^2 + b_3 t^3$

The assumption adopted in a decision to fit form (i) is that the underlying tendency of the series is growth at a constant rate. There are many volume series, in particular index numbers of Production, or of Gross National Product, or other concepts involving a high degree of aggregation, for which the assumption of a constant long run rate of growth is acceptable on theoretical grounds and is also apparent from inspection of the plotted data, particularly if observations are available over half-a-century or longer.

In fitting a trend of form (ii) we assume that the long run rate of growth is increasing or decreasing at a steady rate. For the instantaneous rate of growth is $b_1 + 2b_2t$ and the degree of acceleration or retardation present in the series is constant.[1] Volume series suggesting a long run growth trend subject to constant acceleration are rare. Special factors, for example the adoption of innovations, may serve to account for an accelerated growth trend in specific industries or groups of related industries over several decades [2] but we would not expect this acceleration to be sustained over a longer period. A declining rate of growth, on the other hand, is characteristic of many production and trade series particularly those which are components of a larger aggregate. There is, of course, need for common sense when adopting the assumption that the trend rate of growth increases or decreases with time. The function implied by (ii) is not monotonic. If it is used for purposes of extrapolation the prediction may be unacceptable even though the curve affords a satisfactory picture of the trend within the range for which we have observations.

Constant rates of growth, or rates subject to constant acceleration or retardation, characterise the underlying trend in the case of the majority of series which concern economists. There are, however, occasions when the observations appear to indicate a more complex underlying growth mechanism which can, for certain purposes, be represented by form (iii)—the cubic parabola. In this case the graph of the series on a ratio scale may at first increase at an increasing rate followed by a change to growth at a decreasing rate. Alternatively growth at a decreasing rate may be followed by growth at an increasing rate. In the former case the series will show convex curvature followed by concave (seen from below), in the latter concave followed by convex. The latter possibility is so foreign to our experience of economic or biological growth as to require no

[1]Since this is given by $2b_2$ a negative value for b_2 implies that the rate of growth will decline (and a positive value that it will increase) at a constant rate.

[2]For reasons which should become clear subsequently the charts of deviations from an assumed constant rate of growth of certain series investigated by Lewis and O'Leary do suggest that, over the short period 1870-1913, these series might be considered to be growing at an accelerating rate. Cf. *op. cit.*, p. 117 series 3, p. 133 series 27, 28, 29.

further consideration. The most familiar example of a variable which appears to grow first at an increasing rate and subsequently at a decreasing rate is population. A cubic parabola satisfactorily allows for the successive High Growth Potential, Transitional Growth and Incipient Decline phases suggested by certain demographers. Extrapolation of this function type, however, implies that the Incipient Decline becomes eventually an actual decline. There is no empirical basis for this assumption ; consequently demographers have, in fact preferred to fit a function similar in the central range but which approaches an asymptote rather than turning downwards. This is the logistic

$$y = \frac{k}{1 - be^{-at}}$$

Although the least squares estimate of the constants k and b can only be approximated, the assumptions it implies are more readily acceptable than is the case with the cubic parabola.

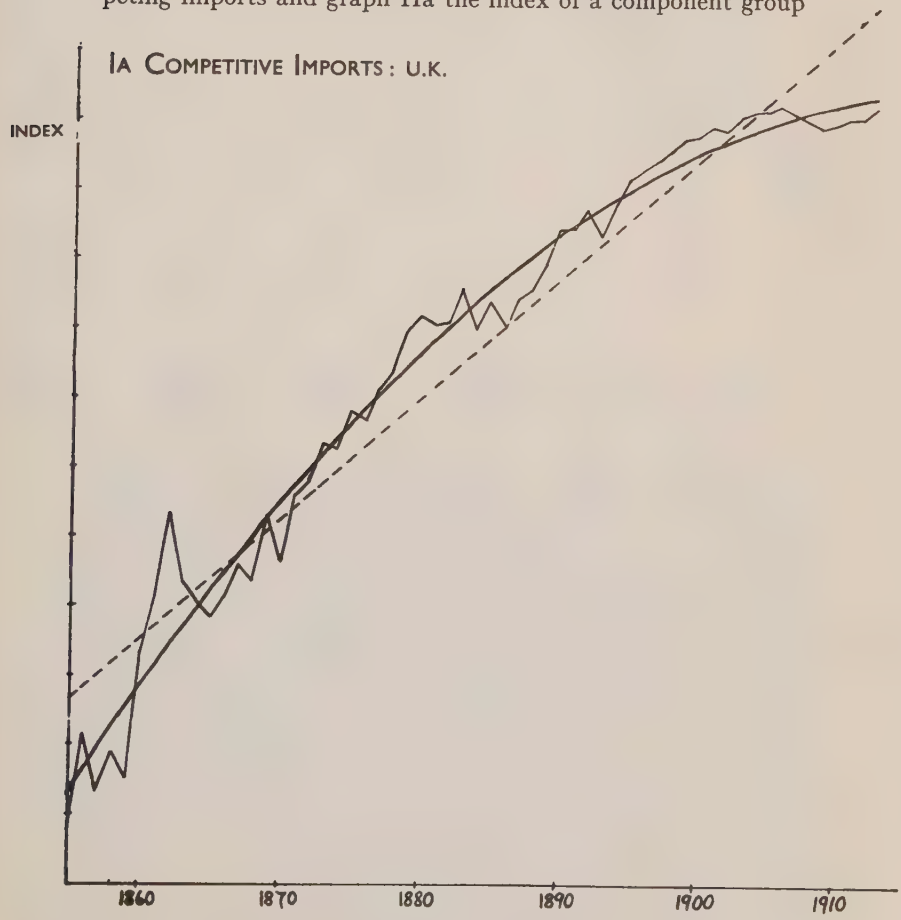
In the majority of situations with which we are confronted the problem in deciding upon an appropriate trend requires a choice between forms (i) and (ii). Whether a straight line or a parabola most accurately describes the growth trend of a series can often be decided merely from inspection of a graph of the series plotted on a ratio scale. If the curvature displayed is pronounced the decision is obvious. In practice a problem arises in situations where the curvature is only slight. The second part of this Note is concerned with the hazards of dependence on subjective criteria in such instances.

What happens to the pattern of deviations if we fit a straight line trend to data which are growing at a declining rate ? The line will pass through the centre of the group of observations. In order to do so it must cut off a segment of the arc traced out by the declining growth trend. It is obvious that the observations will lie below the straight line at the beginning of the period, to an extent depending upon the degree of curvature exhibited by the observations,[1] rising to a

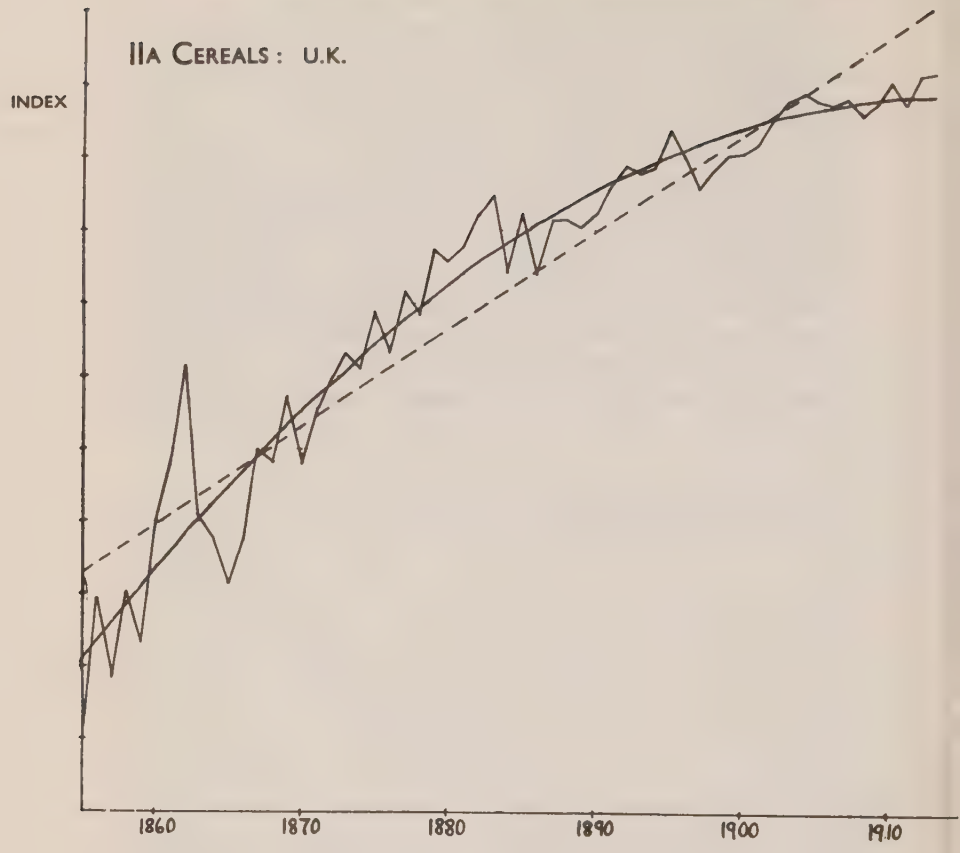
[1] i.e. upon the magnitude of the b_2 coefficient. The examples discussed in this Note illustrate declining growth rates. Similar conclusions apply *mutatis mutandis*, to the consequences of fitting a straight line in the less usual situations in which the series are in fact growing at an increasing rate.

peak above the line somewhere in the central range,[1] and finally falling to reach a value below trend towards the end of the period. In other words the pattern of the deviations will be dominated by one long spurious wave relating the more appropriate parabolic trend values to the inappropriate straight line trend values.

In order to emphasise the distortion of the pattern of deviations which results when we neglect curvature in the series, the first two examples shown in charts I and II serve to contrast the deviations from alternative trends in the case of series displaying a moderate degree of curvature when plotted on a ratio scale. That both these two series grew at a declining rate can be discerned plainly from the graphs IA and IIA. Comparisons of the percentage deviations are shown in the two graphs on the right. Graph Ia shows the index of U.K. competing imports and graph IIa the index of a component group



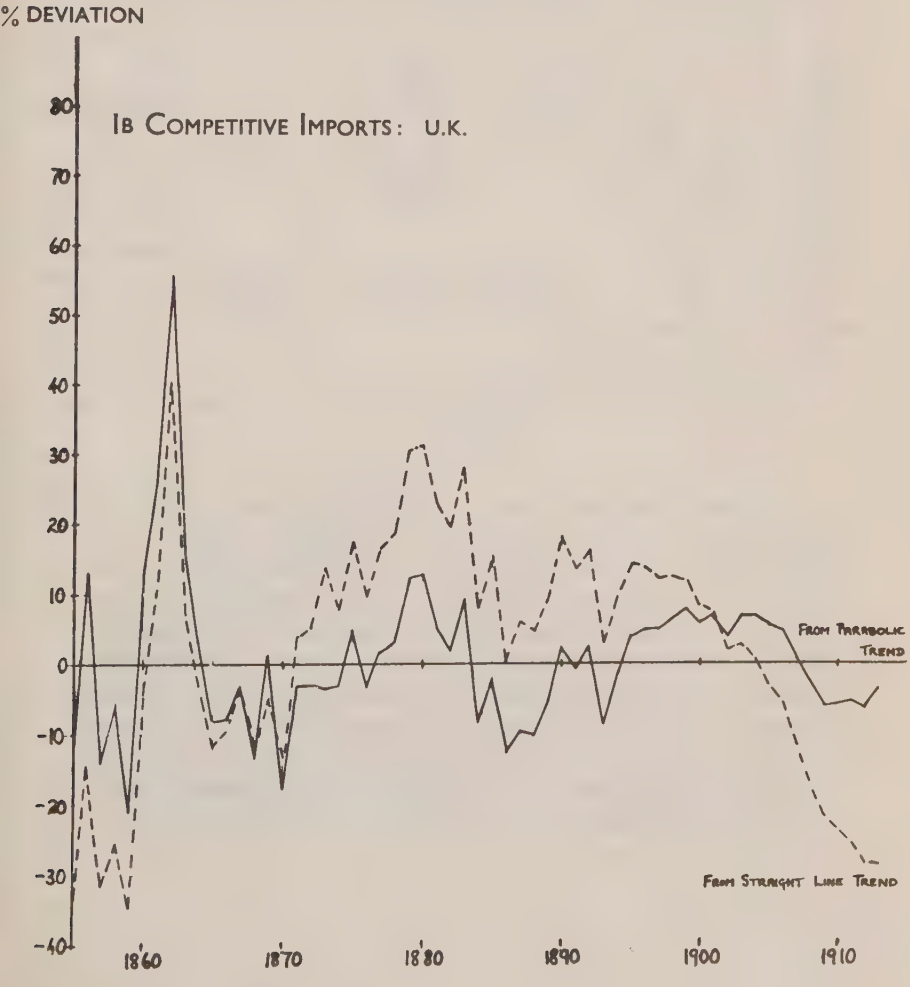
[1]More precisely, the graph of the deviations will tend to have a maximum positive value in the neighbourhood of $t = \frac{-b_1}{2b_2}$. This will not necessarily be the case if the amplitude of the cycle about the trend is large.

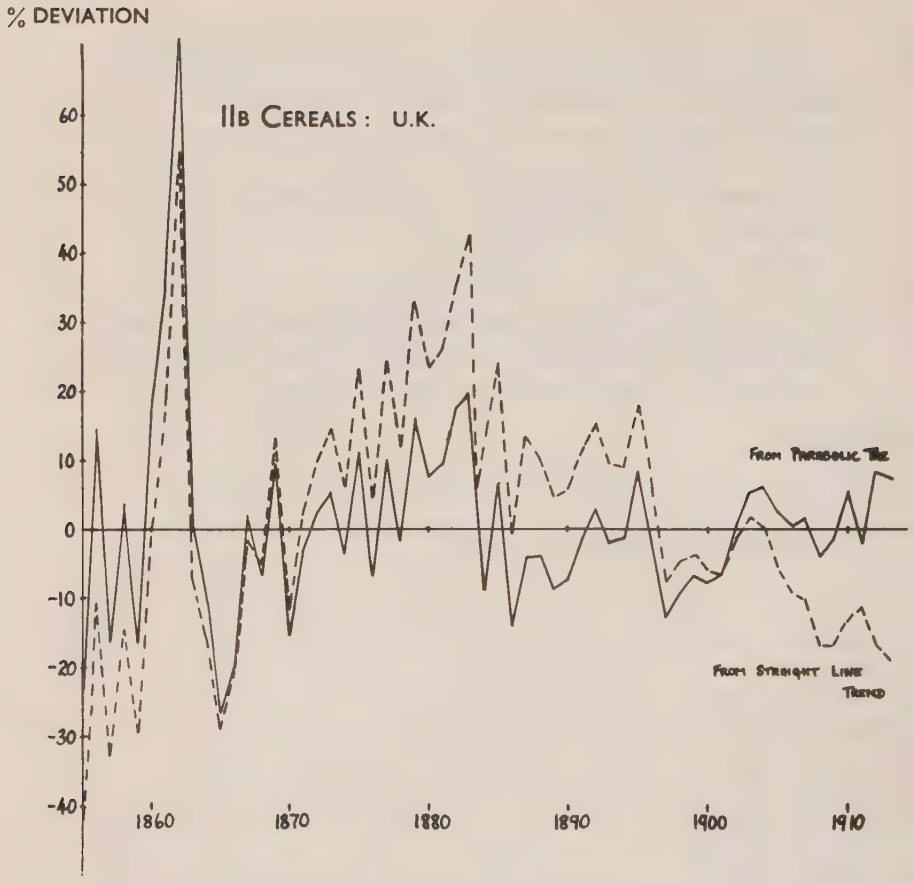


of competing imports :—food cereals. As might be expected the curvature is more evident in the case of the component group. (Over the period 1855–1913 food cereals formed a declining proportion of total competing imports). In both cases the pattern of deviations from a constant growth rate trend gives the illusion of one long secular swing from a trough in the mid-sixties through a peak in the early eighties to a second trough approached just before the first World War. This is eliminated when the deviations are measured from a parabolic trend. Deviations from the latter trend are of greatly reduced amplitude. In the case of all competing imports, by measuring the deviations from a parabolic trend we obtain a secular swing of some twenty years duration with peaks in 1880 and 1899 and troughs in 1870 and 1886. In the case of food cereals we are left with little evidence of any 16–22 year cycle but fairly wide short term fluctuations above and below

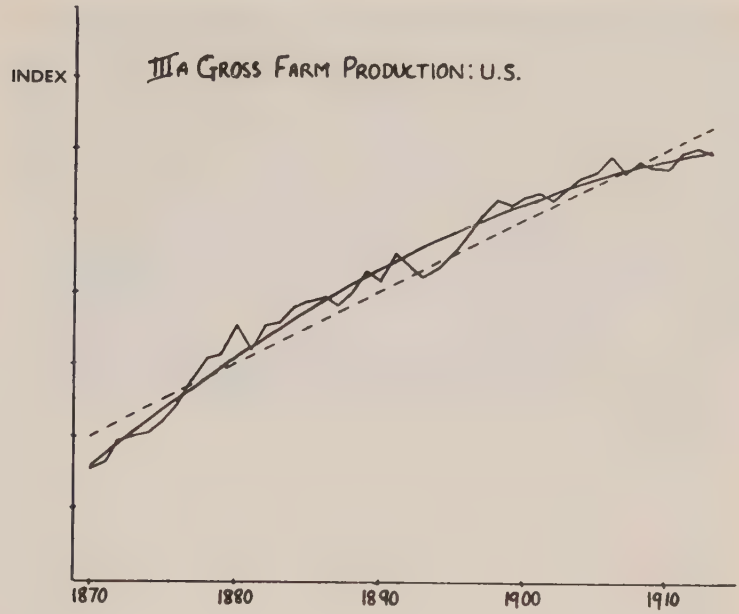
trend. In both these cases we must regard the much wider movement about the straight line trend as the sum of the deviation of an observation from the parabolic trend—which has economic significance—and the deviation of the parabolic trend value from the straight line trend value—which is not only devoid of interpretative significance but also obscures the pattern of cyclical behaviour which we are attempting to discover.

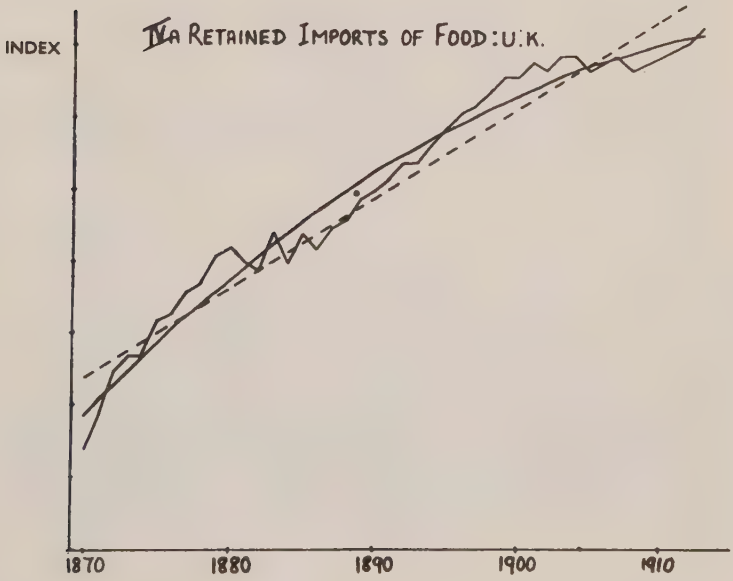
It is unnecessary to dwell further on these first examples. There is little danger of obtaining an impression as grossly misleading as is given by the pattern of deviation from the straight line trend shown in graphs Ib and IIb. Only





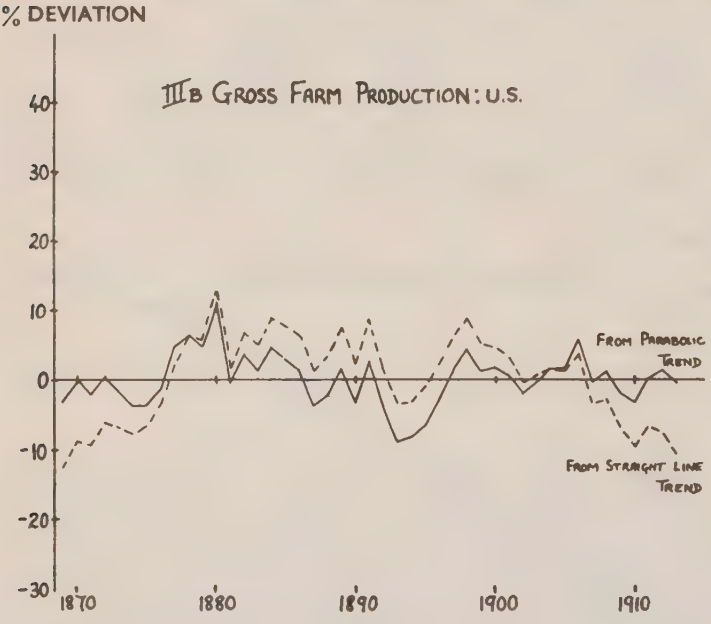
when the departure from a constant growth rate is appreciable is there danger. Inspection of the ratio scale graphs of the original series is quite sufficient to indicate that a straight line trend is inappropriate. We may turn now to see what may happen when the ratio scale graph affords no such unequivocal guidance. Graphs IIIa and IVa show two series investigated by Lewis and O'Leary. From inspection it would not appear that either United States Gross Farm Production or U.K. Retained Imports of Food showed significant departure from a constant rate of growth. With respect to the large number of series they examined Lewis and


INDEX
IIIa Gross Farm Production: U.S.
1870
1880
1890
1900
1910


INDEX
IVa Retained Imports of Food: U.K.
1870
1880
1890
1900
1910

O'Leary point out : "Fortunately most of the series in which we are interested show no retardation between 1870 and 1913, or such retardation that the difference between a first order and a second order parabola is negligible."[1] Charts IIIb and IVb demonstrate that it is desirable to compare the pattern of deviations from the two trends rather than to rely upon what the eye may suggest from inspection of the graphs of the series.

Lewis and O'Leary regard U.S. Farm Production as governed by the joint U.S.-U.K. demand. The United States demand, they say, kept the series above trend until 1891. The U.K. demand produced a "super peak" in 1899 and brought the series "so low" just before the war.[2] Graph IIIb shows that

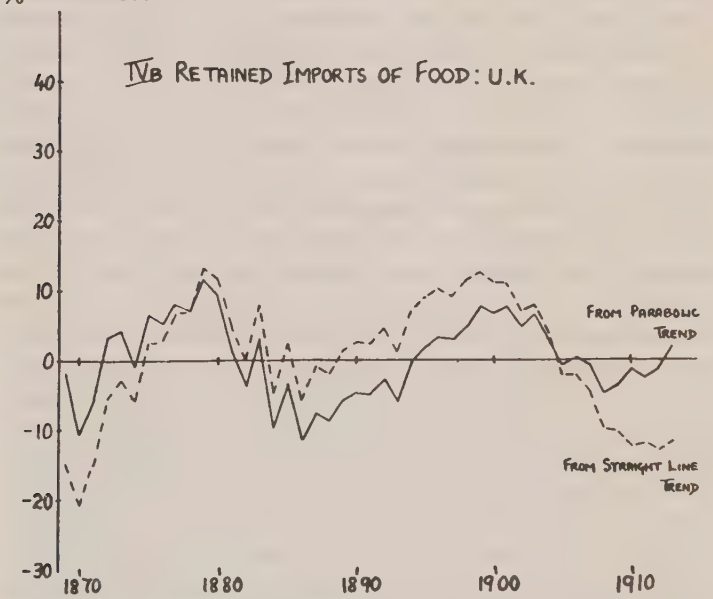


these conclusions cannot be sustained if the deviations around a parabolic trend are considered to measure the cyclical behaviour. The series drops below trend as early as 1887, attains a value in 1898 which cannot, in any sense, be

[1] *Op. cit.*, *p.* 115.
[2] *Ibid.*, p. 138.

considered a "superpeak" and remains close to the trend value in the years immediately before the war. Measured from a parabolic trend there remains some evidence of a secular swing but one so reduced in amplitude as to be scarcely distinguishable from the shorter term movements around the trend. The mean deviation is reduced from 4·1% in the case of the straight line trend to 2·2%.

Similarly graph IVb which compares the deviations from a straight line and from a parabolic trend, in the case of U.K.



Retained Imports of Food,[1] does not leave Lewis and O'Leary's conclusions undisturbed. During the years 1875–1883 and 1894–1906, these authors consider that the U.K. was above trend. They have been rather less precise in indicating years of depression for the U.K. For the world as a whole they consider the Great Depression to cover the years 1883–1896 "but even then it is . . . only of strictly limited application to the U.K." [2]

[1] *Ibid.*, series 19, p. 123.
[2] *Ibid.*, p. 135.

Now, from graph IVb it can be seen that the spurious swing introduced by measuring the deviations from a straight line does not conceal the marked secular swing with peaks in 1879 and 1899 ; but the below trend phases are concentrated at the beginning and end of the period. Between the two above trend phases 1875–1883 and 1894–1906 we do not find depression in the sense of a run of deviations lying below trend. Can it be argued that the Great Depression 1883–1894 is not reflected in U.K. Retained Food Imports ? The deviations from a parabolic trend lie well below trend for the decade 1884–1894 suggesting that this series was, in fact, influenced by World Depression.

Lewis and O'Leary refer to a "rapid rise" in the 'seventies, in spite of declining building, which they attribute to bad harvests, and a decline in the 1900's for which they hold responsible deteriorating terms of trade and a shift in the distribution of income from wages to profits.[1] These influences doubtless operated to some extent but their effect upon retained imports was not as great as Lewis and O'Leary were led to believe by the spurious swing around the straight line trend they had fitted. Retained imports of food do rise in the 'seventies but more slowly than Lewis and O'Leary supposed. There is also some decline in the early 1900's but this is arrested in 1908. Finally the amplitude of the swing around the parabolic trend is less. The mean deviation is 5·5% in the case of the straight line trend and 3·6% in the case of the parabolic trend.

Just how far are we likely to be led astray in basing conclusions upon the pattern of deviations from a straight line trend fitted to a series which is growing at a declining or accelerating rate ? Some indication has been given above of points at which Lewis and O'Leary appear to have reached questionable conclusions but it is necessary to remember that these authors were primarily concerned with the timing of secular swings in relation to one another. Any method of trend elimination which unambiguously defines turning points would lead them to much the same general conclusions. If the residual fluctuation is small relative to the amplitude of the cycle, as in graph IVb, it is easy enough to tell from observation when the series is swinging up or down irrespective of whether the trend

[1] *Ibid.*, p. 124.

we have fitted is a straight line or a parabola. For Lewis and O'Leary's purpose this is for the most part sufficient. In so far, however, as the choice of a parabolic rather than a straight line trend alters the relative significance of turning points, as happens in the case of U.S. Gross Farm Production, ignoring the retardation present in the series may lead to misleading conclusions even when the evidence sought is as limited as that required by Lewis and O'Leary. If in this case these authors had measured the deviations from a parabolic trend it would have been difficult to regard the data as tracing out any defined secular swing.

As soon as we come to demand more from our series, that is when we seek to determine the amplitude of whatever swings may be present, or if interest centres on the timing of below trend or above trend phases, then we can no longer be indifferent to the choice between a straight line and a parabolic trend. It is always possible to reduce the amplitude of a cycle by fitting polynomials of successively higher degree. The more marked the curvature in a series the greater the reduction in amplitude when the pattern of deviations from a parabolic trend is compared with that from a straight line trend.[1]

[1]This, of course, is only another way of saying that, for series containing "true" cycles of given amplitude, if the curvature is appreciable, fitting a second degree polynomial will reduce the scatter of the observations about the trend more than would be the case if the curvature is not appreciable. The competing imports series investigated by the present writer show considerably more retardation than is the case with retained imports (*graph IVa*) partly on account of the greater degree of aggregation in the latter series but mainly on account of the fact that the competing imports series cover the quarter century following the removal of duties on food imports during which imports of food grew much more rapidly than was the case after 1875. As pointed out above the mean deviation in the case of U.K. retained imports of food 1870-1913 was reduced from 5·5 to 3·6—about 35%. In the case of competing imports the comparison is as follows :

Mean Deviations	Total Index	Food Cereals	Meat	Dairy Produce and Lard
From straight line trend	14	14·6	25·8	13·1
From parabolic trend	7·8	9·4	13·6	8·2
% Reduction ..	45	36	47	38

Finally it is evident from the examples shown above that a change of assumption concerning the underlying trend completely alters the balance of the series with respect to above and below trend phases. If we adopt the convention that observations falling above the trend line indicate a period of "prosperity" while those falling below indicate "depression," then we have seen that comparatively small departures from a constant rate of growth can, when we measure the deviations from a straight line trend, result in a spurious swing which provides a run of positive deviations in the middle range of the series balancing negative deviations, at the beginning and the end, which increase in magnitude as the extremes are approached. The misleading conclusions to which this distortion may lead us have been discussed above with reference to U.K. Retained Imports of Food. The table below shows how this alteration in balance happens in the case of all the series examined in this connection by the present writer.

	U.S. Gross Farm Products	U.K. Retained Imports of Food	U.K. Competing Imports			
			Total Index	Food Cereals	Meat	Dairy Produce and Lard
Total number of observations...	45	45	59	59	59	59
Number of negative deviations from straight line becoming positive from parabola ...	4	3	7	11	6	4
Number of positive deviations from straight line becoming negative from parabola ...	6	7	13	13	14	20

In the case of the first two series nearly a quarter of the deviations are changed in sign while in the four competing imports series the proportion is from 33% to 40%. In each of the above examples the majority of changes are from positive to negative and it is easy to see why this is so.

Graphs Ia, IIa, IIIa and IVa show that the majority of the observations falling within the arc formed by the straight line cutting the parabola lie above the straight line, while their distribution around the parabolic trend depends upon the "true" cyclical behaviour, so that within this range, sequences of several years occur during which the observations lie below the parabola and above the straight line.

ROSEMARY WALKER

University of Manchester.

Books Received

ALLEN, E., ODBER, A. J., and BOWDEN, P. J. : *Development Area Policy in the North East of England.* (An inquiry carried out). Report. 5/-. 94 pp.

ALLEN, V. L. : *Trade Union Leadership.* Based on a Study of Arthur Deakin. Longmans, Green & Co. Ltd., London, W.1. 30/- net. 336 pp.

BARAN, P. A. : *The Political Economy of Growth.* Monthly Review Press, N.Y. $5.00. 308 pp.

BUBLOT, Georges : *La Production Agricole Belge Etude Economique Seculaire* 1846-1955. Editions Nauwelaerts, Louvain. 370 fr b. 419 pp.

DATTA, Amlan : *Essays on Economic Development.* Bookland Private Ltd., Calcutta-6. Rs. 5/-. 148 pp.

Economic Issues. A Financial and Economic Debate in the Critical Years, 1954-57. Waterlow and Sons Ltd., London. 21s. 6d. 231 pp.

HARTMAN, Georges : *L'Automation.* Editions de la Baconniere, Boudry, Neuchatel. 242 pp.

ISARD, W. : *Location and Space Economy.* A General Theory Relating to Industrial Location, Market Areas, Land Use, Trade, and Urban Structure. Chapman & Hall, London. 70s. 350 pp.

KRISHNA, Raj, JAIN, L. C., and KRISHNA, Gopi : *Co-operative Farming.* Some Critical Reflections. Indian Co-operative Union. Re. 1/-. 36 pp.

LABINI, Paolo Sylos : *Oligopolio e Progresso Tecnico.* Pubblicazioni della Facolta di Economia E Commercio Dell 'Universita di Roma. 207 pp.

SAILLY, Jean de : *La Zone Sterling.* Librarie Armand Colin, Paris, 5e. 131 pp.

United Nations, N.Y. : *Yearbook of International Trade Statistics* 1955. 720 pp.